

An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points

Charles Sprenger

Stanford University

Recent models of reference-dependent preferences indicate that expectations may play a prominent role in the presence of behavioral anomalies. A subset of such expectations-based models predicts an “endowment effect for risk”: that risk attitudes differ when reference points change from certain to stochastic. In two purposefully simple risk preference experiments, eliminating often-discussed confounds, I demonstrate both between and within subjects such an endowment effect for risk. These results provide needed separation between expectations-based reference-dependent models, allow for evaluation of recent theoretical extensions, and may help to close a long-standing debate in decision science on inconsistency between utility elicitation methodologies.

I. Introduction

Theoretical models of reference-dependent preferences, wherein losses are experienced more severely than commensurate gains, originated in the prospect theory work of Kahneman and Tversky (1979). These models have gained traction, rationalizing experimental anomalies such as the endowment effect and nonstandard findings in a variety of settings including labor market decisions, consumer behavior, and finance, among others.¹

I am grateful for the insightful comments and suggestions of Nageeb Ali, Doug Bernheim, Vince Crawford, David Eil, David Gill, David Laibson, Ted O'Donoghue, Victoria Prowse, Matthew Rabin, and Joel Sobel. Particular thanks are owed to James Andreoni. I also would like to acknowledge the generous support of the National Science Foundation (SES grant 1024683).

¹ The endowment effect refers to the frequent finding in both experimental and survey research that willingness to pay for a given object is generally lower than willingness to

Electronically published October 30, 2015

[*Journal of Political Economy*, 2015, vol. 123, no. 6]

© 2015 by The University of Chicago. All rights reserved. 0022-3808/2015/12306-0004\$10.00

Critical to reference-dependent models is the determination of the reference point around which losses and gains are encoded. Originally, the reference point was left undetermined, taken to be the status quo, current level of assets, or a level of aspiration or expectation (Kahneman and Tversky 1979). Indeed, the freedom of the referent may be the reason why reference dependence is able to rationalize such a large amount of behavior. Model extensions have added needed discipline, with particular attention being given to expectations-based mechanisms for the determination of the referent. An early effort in this vein, disappointment aversion (DA; Bell 1985; Loomes and Sugden 1986; Gul 1991), proposed a fixed referent of the expected utility certainty equivalent of a given prospect.² More recently, Koszegi and Rabin (2006, 2007) (KR) have proposed a referent of the entire distribution of expected outcomes.

The modeling convention of using either the certainty equivalent or the entire distribution of expected outcomes as a referent is more than a minor nuance. The two models make substantively different predictions in a variety of environments from consumer behavior to contracting.³ A central distinction is the existence of a so-called endowment effect for risk, present in the KR model that is absent in DA.⁴ In the KR model, the amount of risk a decision maker is willing to take depends on the riskiness of the referent distribution. In particular, when the referent is a binary gamble and an individual is offered a certain amount on the interior of the gamble's outcome support, the KR model predicts near risk neutrality. Conversely, when the referent is a fixed certain amount

accept for the same good (Thaler 1980; Knetsch and Sinden 1984; Brookshire and Coursey 1987; Coursey, Hovis, and Schulze 1987; Knetsch 1989; Kahneman, Knetsch, and Thaler 1990; Harbaugh, Krause, and Vesterlund 2001). With regard to labor market decisions, reference dependence has proven useful in understanding nonstandard behavior for individuals who make labor market decisions at the daily level (Camerer 1997; Goette and Fehr 2007). For consumer behavior, see, e.g., Hardie, Johnson, and Fader (1993) and Sydnor (2010). Important applications in finance include Odean (1998), Barberis and Huang (2001), and Barberis, Huang, and Santos (2001).

² Disappointment aversion can refer to a number of different classes of models. I focus primarily on those of Bell (1985) and Loomes and Sugden (1986), who capture DA in functional form by fixing the referent as the certainty equivalent or expected value of a given gamble and develop a reference-dependent disappointment-aversion function around this point. Bell (1985) offers some extensions to the referent to incorporate betting on long shots. Shalev (2000) provides a similar functional form in a loss-averse game-theoretic context with the reference point fixed at a gamble's certainty equivalent. Though similar in spirit to these models, Gul (1991) provides a distinct axiomatic foundation for DA relaxing the independence axiom. The resulting representation's functional form is similar to prospect theory probability weighting (Kahneman and Tversky 1979; Tversky and Kahneman 1992), with DA making a particular global restriction on the shape of the probability weighting function (Abdellaoui and Bleichrodt 2007).

³ Section V provides discussion of relevant applications in which the distinction between KR and DA may be important.

⁴ Credit for the term belongs to Koszegi and Rabin (2006, 2007).

and an individual is offered a binary gamble, the KR model predicts risk aversion. Hence, the KR model features an endowment effect for risk in that risk preferences respond to the stochastic structure of the referent.⁵ Disappointment aversion makes no such asymmetric prediction because gambles are always evaluated relative to a fixed referent, the gamble's certainty equivalent.⁶

A recent body of field and laboratory evidence, constructed around plausible manipulation of expectations, has focused on testing the potential importance of expectations-based reference dependence (Post et al. 2008; Abeler et al. 2011; Card and Dahl 2011; Crawford and Meng 2011; Ericson and Fuster 2011; Pope and Schweitzer 2011; Gill and Prowse 2012; Smith 2012; Heffetz and List 2014; Karle, Kirchsteiger, and Peitz 2015; Wenner 2015).⁷ Though both positive and negative results have

⁵ The KR endowment effect for risk should be thought of as distinct from the standard endowment effect literature documenting exchange anomalies such as discrepancies between willingness to accept and willingness to pay for the same object. The KR endowment effect for risk is a risk preference discrepancy between risk taking when the referent is certain and risk taking when the referent is stochastic.

⁶ See Koszegi and Rabin (2007) for further discussion.

⁷ Post et al. (2008) provide evidence from a television game show, *Deal or No Deal*, indicating that whether a contestant accepts or rejects an offer depends on whether he or she has been recently lucky, neutrally lucky, or unlucky in eliminating prizes of low, average, and high value, respectively. The authors propose a model of a path-dependent reference point with something akin to rational expectations adjustment in that the "adjustment is always sufficiently strong to ensure that the reference point is feasible in the next round" (61). Pope and Schweitzer (2011) present primary analysis demonstrating apparent reference dependence in putting performance around par on golf holes, which is complemented by analysis demonstrating responsiveness to average field performance, a plausible expectation. Card and Dahl (2011) document an increase in domestic violence after surprise National Football League losses, indicating a responsiveness to expectations of victory. Crawford and Meng (2011) show that labor market decisions for New York City cabdrivers can be rationalized by a model with expectations-based reference dependence in the domains of hours and earnings. Gill and Prowse (2012) present a two-player real-effort game in which the probability of winning a prize depends linearly on exerted efforts. As the probability dependence is linear, second movers should be unresponsive to first-mover effort. The authors document a discouragement effect consistent with expectations-based reference dependence. In a modification of a standard endowment effect experiment, Ericson and Fuster (2011) demonstrate that individuals with a high experimental probability of being able to exchange are willing to do so with a higher probability than individuals with a low probability of being able to exchange. In contrast, using a very similar experimental design, Heffetz and List (2014) fail to replicate the effects of Ericson and Fuster (2011). Also, Smith (2012) documents a limited relationship between willingness to exchange after the realization of an endowment gamble and the likelihood of endowment. Abeler et al. (2011) present a real-effort experiment in which piece rates are probabilistically replaced with fixed payments. Though the fixed payment should not feature in standard marginal conditions, the authors document responsiveness to the fixed payment, consistent with expectations-based reference dependence. Karle et al. (2015) document a relationship between food choices after the realization of uncertain prices and subsequently measured loss aversion in a manner predicted by expectations-based reference dependence. Using a similar setup with random price realizations, Wenner (2015) fails to find support for expectations-based models.

been documented, an account of the literature would indicate some promise for the relevance of expectations in rationalizing reference-dependent behavior. Importantly, when positive results have been found, the data are generally consistent with either DA or KR. That is, the body of evidence is unable to distinguish between the two models.⁸

The objective of this paper is to fill this gap in the literature by experimentally testing for the endowment effect for risk to distinguish between DA and KR.⁹ Such a distinction is critical for evaluating the theoretical advance provided by KR beyond the prior DA model and for evaluating applications of the two models in settings in which their predictions differ.

Note that the separating prediction of an endowment effect for risk is based on the construction of each model's reference dependence and not on the common force of expectations. Hence, unlike prior studies, an experiment providing such separation need not have manipulation of expectations at its core. However, testing the endowment effect for risk represents a key challenge, as it requires the experimenter to plausibly manipulate the structure of the referent from certain to stochastic across conditions.

Prior exchange experiments provide some suggestion of the critical KR prediction of an endowment effect for risk. Knetsch and Sinden (1984) demonstrate that a higher proportion of individuals are willing to pay \$2 to keep a lottery ticket with unknown odds of winning around \$50 than to accept \$2 to give up the same lottery ticket if they already possess it. Kachelmeier and Shehata (1992) show that selling prices for a 50–50 gamble over \$20 are significantly higher than subsequent buying prices out of experimental earnings for the same gamble. An early debate in decision science identified inconsistencies in risk aversion across utility elicitation methods. In hypothetical decisions, subjects were generally found to be more risk averse when trading a sure amount for a gamble than vice versa (Hershey, Kunreuther, and Schoemaker 1982; Hershey and Schoemaker 1985; McCord and de Neufville 1985, 1986; Schoemaker 1990).¹⁰

Though it may be natural to follow a similar design, such a study could suffer from potential experimental confounds common to the endow-

⁸ For example, in Abeler et al. (2011), the observed responsiveness is consistent with either DA or KR, with no possibility of distinguishing between the two. This point is acknowledged by Abeler et al.

⁹ Recent research outside of this work has also aimed to fill this gap. Song (2012) demonstrates findings similar to those presented here. Castillo and Eil (2014) document a similar endowment effect for risk but also present additional results related to the Allais (1953) common ratio effect, inconsistent with the KR model.

¹⁰ The observation of changing risk posture across elicitation methods, similar to other preference reversals in risky choice (Grether and Plott 1979), called into question the basic tenets of utility theory. See Sec. V for further detail on this debate's objectives and leading opinions.

ment effect literature. Plott and Zeiler (2005, 2007) argue that when endowing subjects with an actual object, such as a lottery ticket, via language, visual cues, or physical cues, subjects may view it as a gift and be unwilling to part with it, thus leading to a demand-induced endowment effect.¹¹ Additionally, for hypothetical risky choice there may be reason for caution as the hypothetical bias has proven to be more than quantitative (Holt and Laury 2002).¹² Given the potential confounds of prior experimental methods, it is important to move away from hypothetical choice, physical endowments, and ownership-related language. Hence, I opt not to follow the prior endowment effect style literature and choose a design without an explicit form of endowment.¹³

I present between- and within-subjects results from simple, neutrally worded risk preference experiments conducted with undergraduate students at the University of California, San Diego. In the primary experiment, 136 subjects were separated into two groups. Half of the subjects were asked a series of certainty equivalents for given binary gambles. In each decision the gamble was fixed while the certain amount was changed. The other half were asked a series of probability equivalents for given certain amounts. In each decision the certain amount was fixed while the binary gamble probabilities were changed. In a second study, a relevant portion of the data collected for Andreoni and Sprenger (2011) with 76 subjects and a similar, within-subjects design are presented.¹⁴ The central objective of these designs is to alter the stochastic structure of the referent across conditions from certain in the probability equivalent tasks to stochastic in the certainty equivalent tasks.¹⁵

¹¹ When using neutral language and elicitation procedures based on the Becker, Degroot, and Marschak (1964) mechanism, Plott and Zeiler (2005) document virtually no difference between willingness to accept and willingness to pay for university-branded mugs. Plott and Zeiler (2007) demonstrate, among other things, the extent to which the endowment effect could be related to subjects' interpretation of gift giving. The authors increase and reduce emphasis on gifts and document corresponding increases and decreases in willingness to trade endowed mugs for pens, and vice versa. Plott and Zeiler (2005) also discuss data from a series of small-scale paid practice lottery conditions, which they argue were contaminated by subject misunderstanding and order effects. Recently these data have been called into question as potentially demonstrating an endowment effect for small-stakes lotteries (Isoni, Loomes, and Sugden 2011). However, the debate remains unresolved as to whether subject misunderstanding of the Becker et al. (1964) mechanism or other aspects of the experimental procedure are the primary factors (Plott and Zeiler 2011).

¹² Additionally, Holt and Laury (2002) note the extent to which the Plott and Zeiler (2005, 2007) confounds may influence risky decisions and "chose to avoid this possible complication by framing the decisions in terms of choices, not purchase and sales" (1644).

¹³ Though my experiment chooses to sidestep the prominent endowment effect concerns, a design with explicit endowments of lotteries and safe options could well prove fruitful. Song (2012) employs such a design and comes to quite similar conclusions.

¹⁴ The data of Andreoni and Sprenger (2011) were collected for an independent project testing the limitations of expected utility's independence axiom.

¹⁵ The chosen design is guided by intuition developed by Koszegi and Rabin (2006), indicating a "first focus" in decision making. More discussion is provided below.

Both between- and within-subjects near risk neutrality is obtained in the certainty equivalents data, while significant risk aversion is obtained in probability equivalents. Subjects randomly assigned to probability equivalent conditions are between three and four times more likely to exhibit risk aversion than subjects assigned to certainty equivalent conditions. These results are maintained when controlling for sociodemographic characteristics, numeracy, cognitive ability, and self-reported risk attitudes. This evidence is consistent with the KR preference model's endowment effect for risk and corresponding KR parameter estimates are close to prior values.

Interpreting the data as supportive of the KR preference model requires that manipulation of the stochastic structure of the referent, and not some other feature of the changing decision environment, delivers the observed treatment effects. A natural criticism for the implemented design is that failing to provide explicit endowments limits the researcher's control of the referent. In order to examine the plausibility that the referent is indeed altered by the experimental design, I explore several pieces of evidence. First, I note that under the assumption of a manipulated referent, the KR preference model rationalizes the data at parameter values close to those previously identified in the literature. Second, I examine behavior in two out-of-sample environments, both choices between binary gambles.¹⁶ These two decision environments bear a striking similarity to each other, but the KR model actually makes dramatically different predictions, provided that the stochastic structure of the referent is manipulated by the design. In the first environment, a KR decision maker should be at times close to risk neutral and at times substantially risk averse. In the second environment, a KR decision maker should always be risk neutral. These unique patterns of behavior are observed in the aggregate, and the previously estimated KR preference model rationalizes the data at standard parameter values.¹⁷ Third, individual-level estimates of the KR model also show a high degree of out-of-sample pre-

¹⁶ The out-of-sample environment is a series of uncertainty equivalents (McCord and de Neufville 1986; Magat, Viscusi, and Huber 1996; Oliver 2005, 2007; Andreoni and Sprenger 2011) conducted in the between-subjects experiment. Uncertainty equivalents ask subjects to choose between a given gamble and alternative gambles outside of the given gamble's outcome support. The KR preference model can make an interesting prediction in this environment. Individuals should be risk averse if the referent is a gamble $(p; y, x)$, $y > x > 0$, and an individual considers gambles $(q; y, 0)$, but should be risk neutral if the referent is a gamble $(p; y, 0)$ and an individual considers gambles $(q; y, x)$. Importantly, standard theories and disappointment aversion again predict no difference in risk preference across this changing experimental environment.

¹⁷ The KR preference model is estimated under the identifying assumption that the experiment manipulates the stochastic structure of the referent. Hence, the out-of-sample predictive validity helps to support this identifying assumption.

dictive validity. Fourth, the KR preference model, under the assumption of an altered referent, outperforms leading alternative explanations in terms of predictive power. Taken together, these findings support the interpretation that the stochastic structure of the referent is plausibly manipulated across conditions and compellingly demonstrate the power of the design.

Interestingly, Koszegi and Rabin (2006) provide intuition for such manipulability in referents, indicating that “a person’s reference point is her probabilistic beliefs about the relevant consumption outcome held between the time she first focused on the decision determining the outcome and shortly before consumption occurs” (1141). This “first-focus” intuition is in line with the psychological literature on “cognitive reference points” (Rosch 1974) and decision anchoring and indicates a sensitivity of referents to experimental variation.¹⁸

In addition to providing techniques for modeling stochastic referents, Koszegi and Rabin (2006, 2007) propose a rational expectations refinement of their model, the preferred personal equilibrium (PPE), in which the referent is revealed by choice behavior. The PPE refinement, implemented at the choice level, predicts identical risk attitudes across the experimental conditions. Since the findings reject disappointment aversion, which predicts the same pattern, they necessarily reject this refinement.¹⁹

Finding evidence of an endowment effect for risk, particularly in a neutral environment like that presented in these studies, provides support for the KR preference model. Complementing prior work demonstrating the importance of expectations for reference dependence, these results are able to distinguish between competing accounts for the structure of the referent across KR and DA. Gaining separation between these models is a necessary experimental step for evaluating theoretical developments that depend critically on the stochasticity of the referent and for a variety of applications in financial decision making, insurance

¹⁸ The first-focus intuition is also in line with evidence from multiperson domains in which behavior is organized around initial reactions to experimental environments (Camerer, Ho, and Chong 2004; Costa-Gomes and Crawford 2006; Crawford and Iriberri 2007; Costa-Gomes, Crawford, and Iriberri 2009).

¹⁹ It may be more appropriate to apply PPE at the level of the experiment as a whole. That is, in equilibrium the referent is a meta-lottery, constructed as the rationally expected distribution of outcomes induced by all decisions in the experiment. Importantly, such a construction may be able to deliver the observed effects of changing risk posture in the between-subjects design, as different individuals face different sets of choices. However, in the within-subjects design, there is no possibility of an experiment-level PPE rationale as each subject faces the same set of choices. Hence, the within-subjects data demonstrating the effect reject the possibility of an experiment-level PPE construction. Online app. A.1 and Sec. IV.B have further detail.

purchasing, and decision science in which the models make differing predictions.

The paper proceeds as follows. Section II presents conceptual considerations for thinking about certainty and probability equivalents in standard theories, reference-dependent theories, and the KR model. Section III presents the experimental design, and Section IV presents results. Section V is a discussion and conclusion.

II. Conceptual Considerations

In this section several models of risk preferences in binary environments are discussed. With one exception, the models predict equivalence of risk attitudes elicited with certainty equivalents and probability equivalents. The exception is the KR model, which predicts an endowment effect for risk.

Consider expected utility. Any complete, transitive, continuous preference ordering over lotteries that also satisfies the independence axiom will be represented by a standard expected utility function, $v(\cdot)$, that is linear in probabilities. Under such preferences, a certainty equivalent for a given gamble will be established by a simple indifference condition. Take a binary p gamble over two positive values, y and $x \leq y$, $(p; y, x)$, and some certain amount, c , satisfying the indifference condition

$$v(c) = p \cdot v(y) + (1 - p) \cdot v(x).$$

Under expected utility, whether risk preferences are elicited via the certainty equivalent, c , or the probability equivalent, p , the elicited level of risk aversion, or the curvature of $v(\cdot)$, should be identical. The conformity of risk preference across methodologies is what I term having no endowment effect for risk.

A similar argument can be made for cumulative prospect theory, which establishes loss-averse utility relative to some fixed referent and relaxes the independence axiom's implied linearity in probability (Tversky and Kahneman 1992; Tversky and Fox 1995; Wu and Gonzalez 1996; Prelec 1998; Gonzalez and Wu 1999; Abdellaoui 2000; Bleichrodt and Pinto 2000). Let $u(\cdot | r)$ represent loss-averse utility given some fixed referent, r . The cumulative prospect theory indifference condition is

$$u(c | r) = \pi(p) \cdot u(y | r) + [1 - \pi(p)] \cdot u(x | r),$$

where $\pi(\cdot)$ represents some arbitrary nonlinear probability weighting function. Under such a utility formulation, certainty and probability equivalents again yield identical risk attitudes as the reference point is fixed at some known value.

Extensions to reference-dependent preferences have attempted to explain behavior by establishing what the reference point should actually be. Important for this paper are DA models, which fix the prospect theory reference point via expectations as a gamble's expected utility certainty equivalent (Bell 1985; Loomes and Sugden 1986). DA's fixed referent does not change the predicted equivalence of risk preferences across probability and certainty equivalents as gambles are always evaluated relative to their certainty equivalents. In effect, DA selects the referent for the gamble in the above indifference condition as the expected utility certainty equivalent, $v^{-1}(p \cdot v(y) + (1 - p) \cdot v(x))$, and selects the referent for the certain amount to be c itself.²⁰

A. KR Preferences

The KR model builds on standard reference-dependent preferences in two important ways. First, similarly to DA, the referent is expectations based. Second, and important for this study, the referent may be stochastic. This second innovation implies that behavior when the referent is stochastic will be substantially different than when the referent is certain. In particular, KR preferences as presented below predict risk neutrality in specific cases in which the referent is stochastic and risk aversion in cases in which the referent is certain.

Let r represent the referent drawn according to measure G . Let x be a consumption outcome drawn according to measure F . Then the KR utility formulation is

$$U(F | G) = \iint u(x | r) dG(r) dF(x)$$

with

$$u(x | r) = m(x) + \mu(m(x) - m(r)).$$

²⁰ Take, e.g., a linear $v(\cdot)$ function and a linear $\pi(\cdot)$ function such that risk posture is delivered by reference dependence alone. The gamble's referent is thus $p \cdot y + (1 - p) \cdot x$. Under the reference-dependent utility formulation described below and summarized by loss aversion parameter $\lambda > 1$, the utility of the gamble is

$$p(y + \{y - [p \cdot y + (1 - p) \cdot x]\}) + (1 - p)(x + \lambda\{x - [p \cdot y + (1 - p) \cdot x]\})$$

or

$$p \cdot y + (1 - p) \cdot x + p(1 - p) \cdot (y - x) \cdot (1 - \lambda),$$

which is expected value penalized by loss aversion. The indifference condition is thus

$$c = p \cdot y + (1 - p) \cdot x + p(1 - p) \cdot (y - x) \cdot (1 - \lambda),$$

and the extent of risk aversion in DA, here summarized by λ , can be recovered via probability or certainty equivalents as both will deliver the same behavior.

The function $m(\cdot)$ represents consumption utility and $\mu(\cdot)$ represents gain-loss utility relative to the referent, r . Several simplifying assumptions are made. First, following Koszegi and Rabin (2006, 2007), small-stakes decisions are considered such that consumption utility, $m(\cdot)$, can plausibly be taken as approximately linear, and a piecewise-linear gain-loss utility function is adopted,

$$\mu(y) = \begin{cases} \eta \cdot y & \text{if } y \geq 0 \\ \eta \cdot \lambda \cdot y & \text{if } y < 0, \end{cases}$$

where the utility parameter λ represents the degree of loss aversion. For simplicity and to aid the exposition, $\eta = 1$ is assumed, and only binary lotteries are considered such that G and F will be summarized by probabilities p and q , respectively.

Consider two cases: first where the referent is certain and consumption outcomes are stochastic and, second, where the referent is stochastic and consumption outcomes are certain. The above KR model predicts risk-averse behavior in the first case and risk neutrality in the second. This is illustrated next.

B. Probability Equivalent: Certain Referent, Binary Consumption Gamble

Consider a referent, r , and a binary consumption gamble with outcomes $x_1 \geq r$ with probability q and $x_2 \leq r$ with probability $1 - q$.²¹ Write the KR utility as

$$U(F | r) = q \cdot u(x_1 | r) + (1 - q) \cdot u(x_2 | r). \quad (1)$$

The first term refers to the chance of obtaining x_1 as the consumption outcome with r as the referent. The second term is similar for obtaining x_2 . If $x_1 \geq r > x_2$, the KR model predicts loss aversion to be present in the second term. Under the assumptions above, this becomes

$$U(F | r) = q \cdot [x_1 + 1 \cdot (x_1 - r)] + (1 - q) \cdot [x_2 + \lambda \cdot (x_2 - r)]. \quad (2)$$

Compare this to the utility of the certain amount, $U(r|r) = r$. The lottery will be preferred to the certain referent if $U(F|r) > U(r|r)$ and the indifference point, or *probability equivalent*, will be obtained for some F^* , with corresponding probability q^* , such that $U(F^*|r) = U(r|r)$:

$$r = q^* \cdot [x_1 + 1 \cdot (x_1 - r)] + (1 - q^*) \cdot [x_2 + \lambda \cdot (x_2 - r)],$$

$$q^* = \frac{r - x_2 - \lambda \cdot (x_2 - r)}{(x_1 - x_2) + [1 \cdot (x_1 - r) - \lambda \cdot (x_2 - r)]}. \quad (3)$$

²¹ I assume that $x_2 \geq 0$ and that at least one of the inequalities is strict such that the consumption gamble is nondegenerate.

Note, for an individual who is not loss averse, $\lambda = 1$, $q^* = (r - x_2)/(x_1 - x_2)$, the risk-neutral level. For loss-averse individuals with $\lambda > 1$, $q^* > (r - x_2)/(x_1 - x_2)$ for $x_1 > r > x_2 \geq 0$. The gamble F^* will have higher expected value than r . Figure 1, panel A, illustrates the decision for a loss-averse individual. Additionally, $dq^*/d\lambda > 0$ for $x_1 > r > x_2 \geq 0$, such that probability equivalents are increasing in the degree of loss aversion.²² When the referent is a fixed amount and an individual considers a probability equivalent binary gamble, a loss-averse individual will appear risk averse.

C. Certainty Equivalent: Binary Referent Gamble, Certain Consumption

Now consider a binary referent gamble and the prospect of certain consumption. Let r_1 be the referent with probability p and $r_2 \leq r_1$ be the referent with probability $1 - p$. The utility of the binary referent gamble is

$$U(G | G) = p \cdot p \cdot u(r_1 | r_1) + (1 - p) \cdot (1 - p) \cdot u(r_2 | r_2) \\ + p \cdot (1 - p) \cdot u(r_1 | r_2) + p \cdot (1 - p) \cdot u(r_2 | r_1).$$

The first term refers to the chance of obtaining r_1 as the consumption outcome with r_1 as the referent. The second term is similar for r_2 . The third term refers to the chance of obtaining r_1 as the consumption outcome with r_2 as the referent. The fourth term refers to the chance of obtaining r_2 as the consumption outcome with r_1 as the referent. With $r_1 \geq r_2$, the KR model predicts loss aversion to be present in the fourth term. Under the assumed utility formulation this reduces to

$$U(G | G) = p^2 \cdot r_1 + (1 - p)^2 \cdot r_2 + p \cdot (1 - p) \cdot [r_1 + 1 \cdot (r_1 - r_2)] \\ + p \cdot (1 - p) \cdot [r_2 + \lambda \cdot (r_2 - r_1)], \\ U(G | G) = p \cdot r_1 + (1 - p) \cdot r_2 + p \cdot (1 - p) \\ \cdot [1 \cdot (r_1 - r_2) + \lambda \cdot (r_2 - r_1)].$$

Given this stochastic referent, consider the utility of a certain outcome, x , with $r_1 \geq x \geq r_2$,

$$U(x | G) = p \cdot u(x | r_1) + (1 - p) \cdot u(x | r_2), \\ U(x | G) = x + p \cdot [\lambda \cdot (x - r_1)] + (1 - p) \cdot [1 \cdot (x - r_2)].$$

²² The derivative

$$dq^*/d\lambda = \frac{-(x_2 - r) \cdot (2x_1 - 2r)}{[x_1 - x_2 + 1 \cdot (x_1 - r) - \lambda \cdot (x_2 - r)]^2} > 0$$

for $x_1 > r > x_2 \geq 0$. Note that as x_1 approaches r then q^* approaches one; and as x_2 approaches r then q^* approaches zero. Hence q^* will accord with the risk-neutral level, $(r - x_2)/(x_1 - x_2)$, at the limits $x_1 = r$ and $x_2 = r$. This implies a hump-shaped deviation between q^* and the risk-neutral level of q if $\lambda > 1$.

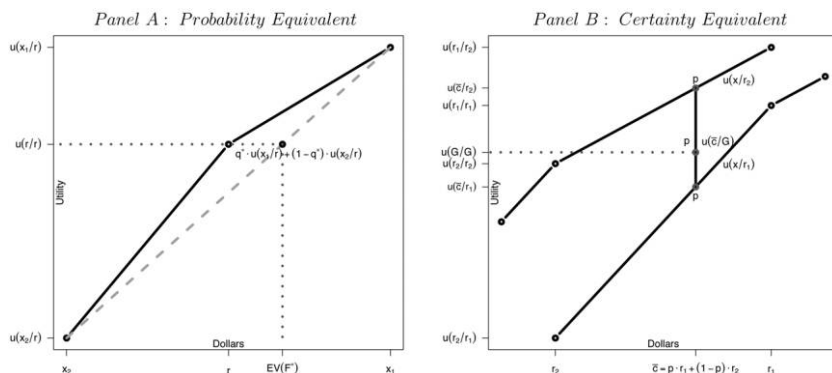


FIG. 1.—KR probability and certainty equivalents. The figure illustrates probability and certainty equivalents under KR preferences. For probability equivalents in panel A, the KR model predicts apparent risk-averse behavior as the expected value of the probability equivalent gamble, $EV(F^*)$, is greater than the referent, r . For certainty equivalents in panel B, KR predicts risk neutrality at all levels of loss aversion as the utility of the expected value of a gamble, $u(\bar{c}|G)$, is equal to the value of the gamble, $u(G|G)$.

The indifference point, or *certainty equivalent*, c , is obtained for $U(c | G) = u(G | G)$:

$$p \cdot r_1 + (1-p) \cdot r_2 + p \cdot (1-p) \cdot [1 \cdot (r_1 - r_2) + \lambda \cdot (r_2 - r_1)] \\ = c + p \cdot [\lambda \cdot (c - r_1)] + (1-p) \cdot [1 \cdot (c - r_2)],$$

To demonstrate that individuals will be risk neutral in certainty equivalent decisions, one need only establish the expected value as the risk-neutral benchmark, $\bar{c} = p \cdot r_1 + (1-p) \cdot r_2$. Substituting $c = \bar{c}$ in the right-hand side of the above equation, one obtains

$$p \cdot r_1 + (1-p) \cdot r_2 + p \cdot [\lambda(p \cdot r_1 + (1-p) \cdot r_2 - r_1)] \\ + (1-p) \cdot \{1[p \cdot r_1 + (1-p) \cdot r_2 - r_2]\},$$

which reduces to

$$p \cdot r_1 + (1-p) \cdot r_2 + p \cdot (1-p) \cdot [1(r_1 - r_2) + \lambda(r_2 - r_1)]$$

and is identical to the left-hand side of the above equation. Hence, indifference occurs at the risk-neutral benchmark, $c = \bar{c} = p \cdot r_1 + (1-p) \cdot r_2$. Figure 1, panel B, illustrates the logic geometrically. In a certainty equivalent, the evaluation of the gamble's expected value never necessitates evaluating over a loss-averse kink point. When the referent is a binary gamble and an individual considers a certainty equivalent inside the gamble's outcome support, a loss-averse individual will appear risk neutral, regardless of the level of loss aversion. This is in contrast to probability equivalents

in which loss-averse individuals appear risk averse. It is this distinction between certainty equivalents and probability equivalents that reflects the endowment effect for risk.

Schmidt, Starmer, and Sugden (2008) discuss a next generation of prospect-theoretic preferences, third-generation prospect theory (or PT³), which combines the basic structure of the KR stochastic referent with cumulative prospect theory probability distortions. The authors document that PT³ is consistent with a valuation gap for risky gambles not predicted under standard prospect theory. Hence, the effects documented in the present suite of experiments are consistent with either KR or PT³. One critical difference between the two models that may provide for future experimental distinction is that in the KR model, states are evaluated as if they were statistically independent while in PT³, reference dependence is state contingent.

D. Rational Expectations Equilibrium Behavior

Koszegi and Rabin (2006, 2007) present a rational expectations equilibrium concept, the unacclimating personal equilibrium (UPE). The objective of the UPE concept is to represent the notion that rational individuals will expect as the referent only outcomes that they will definitely consume given the expectation of said outcomes. Hence, consumption outcomes and referents coincide in UPE. To select among the potential multiplicity of such equilibria, the KR model features a refinement, the PPE. The PPE concept maintains that the UPE with the highest ex ante expected utility is selected.²³

The development above demonstrates that KR preferences may allow for a difference in elicited risk behavior between certainty equivalents and probability equivalents. However, this difference is not predicted under PPE at the choice level. The probability equivalent, $U(F^*|r)$, and the certainty equivalent, $U(c|G)$, are not UPE values as the referent and consumption outcomes do not coincide.

If the referent is revealed in choice behavior, then when an individual is observed accepting some gamble, F^* , over some fixed amount, r , the PPE concept establishes only that $U(F^*|F^*) > U(r|r)$, provided that both $U(F^*|F^*)$ and $U(r|r)$ are UPE. That is, $(F^*|F^*)$ provided the higher ex ante expected utility. If $U(F^*|F^*) > U(r|r)$ is the PPE revealed preference in a probability equivalent, then it cannot be that the opposite is revealed in

²³ Another equilibrium concept in Koszegi and Rabin (2006, 2007) is the choice-acclimating personal equilibrium (CPE), which applies to decisions made far in advance of the resolution of uncertainty. The operational distinction between the two concepts is that a CPE need not be UPE, but a PPE must be UPE. In the present context, CPE and PPE have similar implications, as both are based on the coincidence of referent and consumption outcomes.

a certainty equivalent. Under PPE applied at the choice level, the KR model predicts no difference between certainty equivalents and probability equivalents.

Two further considerations are made. First, rational expectations equilibrium behavior at the choice level may be an inappropriate degree of granularity. In many risky choice experiments (including those presented here) individuals are given a series of decisions, with one chosen randomly for implementation.²⁴ A referent set in rational expectations is plausibly a meta-lottery, constructed as the rationally expected distribution of outcomes induced by all decisions in the experiment. Interestingly, with such a construction one may uncover a PPE endowment effect for risk in a between-subjects design if different subjects face different choice sets. Online appendix A.1 provides detail.²⁵ However, in a within-subjects experiment, where individuals face both certainty equivalents and probability equivalents, the endowment effect for risk should not be present in PPE. The construction cannot generate the preference reversals required for the effect at the individual level. Online appendix A.1 provides a discussion. Hence, the within-subjects data presented here may helpfully distinguish this more broadly bracketed notion of PPE.

Second, equilibrium itself may be a challenging requirement. Individuals may have their referent changed in accordance with changes in contextual variables. Koszegi and Rabin (2006) provide intuitive support for such a possibility suggesting that the referent is established at the moment of first focus on a decision. If first focus is drawn to a specific decision element, then manipulation of this element can induce changes of the referent. The design is indeed predicated on the first-focus notion that experimental manipulation can alter the structure of the referent. To induce a certain referent, individuals face a series of decisions between an initially presented fixed certain outcome and changing gambles. To induce a stochastic referent, individuals face a series of decisions between an initially presented fixed gamble, which was presented first, and changing certain outcomes. The researcher's potential inability to induce changes to the referent with such a manipulation should be viewed as a limitation of the proposed experimental design.

²⁴ See Sec. III for design detail and further discussion.

²⁵ For ease of explication, the examples considered in online app. A.1 focus on small, three-decision choice sets. For three-decision choice sets, 16 utility terms must be calculated. Applying the PPE concept to the body of experimental data collected here would necessitate calculation of thousands of utility terms, though the logic is identical. The intuition is that in a probability equivalent, subjects may be more risk averse in equilibrium as they can rationally expect a certain outcome with high probability and fulfill this expectation. Certainty equivalents carry no such promise as choosing changing certain outcomes induces uncertainty regardless. See online app. A.1 for further detail.

III. Experimental Design

Motivated by the conceptual development above, a primary between-subjects two-condition experiment was designed. A secondary within-subjects design with similar methods and data from Andreoni and Sprenger (2011) is discussed in Section IV.B. In condition 1, subjects completed two series of probability equivalent tasks. The tasks were designed in price list style with 21 decision rows in each task. Each decision row was a choice between option A, a certain amount, and option B, an uncertain gamble. The certain option A was fixed for each task, as were the gamble outcomes. The probability of receiving the gamble's good outcome increased from 0 percent to 100 percent as subjects proceeded through the task. In condition 1.1, subjects completed eight tasks with fixed certain amounts chosen from {\$6, \$8, \$10, \$14, \$17, \$20, \$23, \$26} and gambles over \$30 and \$0. In condition 1.2, subjects completed six tasks with fixed certain amounts chosen from {\$12, \$14, \$17, \$20, \$23, \$26} and gambles over \$30 and \$10. Most subjects began each task by preferring option A and then switched to option B such that the probability at which a subject switches from option A to option B provides bounds for his or her probability equivalent. Panel A of table 1 features a sample probability equivalent task.²⁶ If the fixed option A certain amount in each task is perceived as the referent, the KR model predicts risk aversion in these probability equivalents.

In condition 2, subjects completed two series of certainty equivalent tasks. The tasks were similarly designed in price list style with 22 decision rows in each task. Each row was a choice between option A, a gamble, and option B, a certain amount. The option A gamble was fixed for each task. The certain amount increased as subjects proceeded down the task. In condition 2.1, subjects completed seven tasks with gamble outcomes of \$30 and \$0, probabilities chosen from {.05, .10, .25, .50, .75, .90, .95}, and certain amounts ranging from \$0 to \$30. In condition 2.2, subjects completed a further seven tasks with gamble outcomes of \$30 and \$10, probabilities chosen from {.05, .10, .25, .50, .75, .90, .95}, and certain amounts ranging from \$10 to \$30. Most subjects began each task by preferring option A and then switched to option B such that the certain value at which a subject switched from option A to option B provides bounds for their certainty equivalent. Panel B of table 1 features a sample certainty equivalent task. If the fixed option A gamble in each task is perceived as the referent, the KR model predicts risk neutrality in these certainty equivalents.

A. Additional Measures

The probability and certainty equivalent tasks of conditions 1.1, 1.2, 2.1, and 2.2 provide a simple comparison of elicited risk attitudes. This

²⁶ Online app. A.5 provides the full set of instructions and tasks.

TABLE 1
PROBABILITY AND CERTAINTY EQUIVALENTS: CONDITIONS 1.1 AND 2.1

A. PROBABILITY EQUIVALENT				B. CERTAINTY EQUIVALENT			
Option A		Option B		Option A		Option B	
Certain Payment of \$10		Chance of \$0	Chance of \$30	Chance of \$0	Chance of \$30	Certain Amount	
1. \$10	<input type="checkbox"/>	100 in 100	<input type="checkbox"/>	1. 50 in 100	<input type="checkbox"/>	\$0.00 with certainty	<input type="checkbox"/>
2. \$10	<input type="checkbox"/>	95 in 100	5 in 100	2. 50 in 100	<input type="checkbox"/>	\$0.50 with certainty	<input type="checkbox"/>
3. \$10	<input type="checkbox"/>	90 in 100	10 in 100	3. 50 in 100	<input type="checkbox"/>	\$1.00 with certainty	<input type="checkbox"/>
4. \$10	<input type="checkbox"/>	85 in 100	15 in 100	4. 50 in 100	<input type="checkbox"/>	\$1.50 with certainty	<input type="checkbox"/>
5. \$10	<input type="checkbox"/>	80 in 100	20 in 100	5. 50 in 100	<input type="checkbox"/>	\$2.50 with certainty	<input type="checkbox"/>
6. \$10	<input type="checkbox"/>	75 in 100	25 in 100	6. 50 in 100	<input type="checkbox"/>	\$3.50 with certainty	<input type="checkbox"/>
7. \$10	<input type="checkbox"/>	70 in 100	30 in 100	7. 50 in 100	<input type="checkbox"/>	\$4.50 with certainty	<input type="checkbox"/>
8. \$10	<input type="checkbox"/>	65 in 100	35 in 100	8. 50 in 100	<input type="checkbox"/>	\$6.50 with certainty	<input type="checkbox"/>
9. \$10	<input type="checkbox"/>	60 in 100	40 in 100	9. 50 in 100	<input type="checkbox"/>	\$8.50 with certainty	<input type="checkbox"/>
10. \$10	<input type="checkbox"/>	55 in 100	45 in 100	10. 50 in 100	<input type="checkbox"/>	\$10.50 with certainty	<input type="checkbox"/>
11. \$10	<input type="checkbox"/>	50 in 100	50 in 100	11. 50 in 100	<input type="checkbox"/>	\$13.50 with certainty	<input type="checkbox"/>
12. \$10	<input type="checkbox"/>	45 in 100	55 in 100	12. 50 in 100	<input type="checkbox"/>	\$16.50 with certainty	<input type="checkbox"/>
13. \$10	<input type="checkbox"/>	40 in 100	60 in 100	13. 50 in 100	<input type="checkbox"/>	\$19.50 with certainty	<input type="checkbox"/>
14. \$10	<input type="checkbox"/>	35 in 100	65 in 100	14. 50 in 100	<input type="checkbox"/>	\$21.50 with certainty	<input type="checkbox"/>
15. \$10	<input type="checkbox"/>	30 in 100	70 in 100	15. 50 in 100	<input type="checkbox"/>	\$23.50 with certainty	<input type="checkbox"/>
16. \$10	<input type="checkbox"/>	25 in 100	75 in 100	16. 50 in 100	<input type="checkbox"/>	\$25.50 with certainty	<input type="checkbox"/>
17. \$10	<input type="checkbox"/>	20 in 100	80 in 100	17. 50 in 100	<input type="checkbox"/>	\$26.50 with certainty	<input type="checkbox"/>
18. \$10	<input type="checkbox"/>	15 in 100	85 in 100	18. 50 in 100	<input type="checkbox"/>	\$27.50 with certainty	<input type="checkbox"/>
19. \$10	<input type="checkbox"/>	10 in 100	90 in 100	19. 50 in 100	<input type="checkbox"/>	\$28.50 with certainty	<input type="checkbox"/>
20. \$10	<input type="checkbox"/>	5 in 100	95 in 100	20. 50 in 100	<input type="checkbox"/>	\$29.00 with certainty	<input type="checkbox"/>
21. \$10	<input type="checkbox"/>	0 in 100	100 in 100	21. 50 in 100	<input type="checkbox"/>	\$29.50 with certainty	<input type="checkbox"/>
				22. 50 in 100	<input type="checkbox"/>	\$30.00 with certainty	<input type="checkbox"/>

design is complemented with a third set of tasks for which the KR model can also predict experimental differences, which again are distinct from the predictions of DA. This third set of tasks also serves as an opportunity to investigate the plausibility that the experimental manipulation does indeed alter the stochastic structure of the referent.

Condition 1.3, completed by subjects assigned to condition 1, was a series of eight uncertainty equivalent tasks with 21 decision rows in each task. Option A was a fixed gamble over \$30 and \$10 with probabilities chosen from $\{.00, .05, .10, .25, .50, .75, .90, .95\}$. Option B was a changing gamble over \$30 and \$0. Condition 2.3, completed by subjects assigned to condition 2, was a series of inverted uncertainty equivalent tasks with 21 decision rows in each task. Option A was a fixed gamble over \$30 and \$0 with probabilities chosen from $\{.35, .40, .50, .60, .75, .85, .90, .95\}$. Option B was a changing gamble over \$30 and \$10. Table 2 (panels A and B) provides sample tasks.

The KR preference model can predict a marked difference in elicited risk attitudes across conditions 1.3 and 2.3 despite their apparent similarity. If the referent is perceived as the fixed element in each task, the KR model predicts a unique shape of quadratically declining risk aversion in condition 1.3, the standard uncertainty equivalent. The reason is that at the lowest probability, 0, the task is identical to a probability equivalent of \$10 for sure. As discussed in Section II.B, risk aversion is predicted. At higher probabilities both the referent and the outcomes are stochastic, leading to diminished risk aversion. At probabilities close to 1, near risk neutrality is predicted. The uncertainty equivalent ($q; 30, 0$) for a given gamble ($p; 30, 10$) will be a convex function of p related to the squared probability, p^2 . This quadratically declining deviation from risk neutrality depends on the degree of loss aversion, λ . Online appendix section A.2 provides the detail.²⁷

If the referent is perceived as the fixed decision element in the inverted uncertainty equivalents of condition 2.3, the KR model predicts risk neutrality despite the tasks bearing a striking resemblance to those of condition 1.3. Having prospective gamble outcomes inside of the support of the referent gamble is similar to having a perturbed certainty equivalent task. Just as risk neutrality is predicted in certainty equivalents, the KR model also predicts risk neutrality in the inverted uncertainty equivalents as utility is never evaluated over a loss-averse kink point. Online appendix section A.2 again provides the mathematical detail. Conditions 1.3 and 2.3 also separate the KR and DA models. Though DA can deliver a quadratically declining deviation from risk

²⁷ This nonlinear prediction is in contrast to the prediction of expected utility in which q should be a linear function of p . See Andreoni and Sprenger (2011) for a discussion of uncertainty equivalents and their value in separating between competing models of risk preferences.

TABLE 2
UNCERTAINTY EQUIVALENTS: CONDITIONS 1.3 AND 2.3

A. UNCERTAINTY EQUIVALENT				B. INVERTED UNCERTAINTY EQUIVALENT			
Option A		Option B		Option A		Option B	
Chance of \$10	Chance of \$30	Chance of \$0	Chance of \$30	Chance of \$0	Chance of \$30	Chance of \$10	Chance of \$30
1. 50 in 100	50 in 100	or	0 in 100	1. 50 in 100	50 in 100	or	0 in 100
2. 50 in 100	50 in 100	or	5 in 100	2. 50 in 100	50 in 100	or	5 in 100
3. 50 in 100	50 in 100	or	10 in 100	3. 50 in 100	50 in 100	or	10 in 100
4. 50 in 100	50 in 100	or	15 in 100	4. 50 in 100	50 in 100	or	15 in 100
5. 50 in 100	50 in 100	or	20 in 100	5. 50 in 100	50 in 100	or	20 in 100
6. 50 in 100	50 in 100	or	25 in 100	6. 50 in 100	50 in 100	or	25 in 100
7. 50 in 100	50 in 100	or	30 in 100	7. 50 in 100	50 in 100	or	30 in 100
8. 50 in 100	50 in 100	or	35 in 100	8. 50 in 100	50 in 100	or	35 in 100
9. 50 in 100	50 in 100	or	40 in 100	9. 50 in 100	50 in 100	or	40 in 100
10. 50 in 100	50 in 100	or	45 in 100	10. 50 in 100	50 in 100	or	45 in 100
11. 50 in 100	50 in 100	or	50 in 100	11. 50 in 100	50 in 100	or	50 in 100
12. 50 in 100	50 in 100	or	55 in 100	12. 50 in 100	50 in 100	or	55 in 100
13. 50 in 100	50 in 100	or	60 in 100	13. 50 in 100	50 in 100	or	60 in 100
14. 50 in 100	50 in 100	or	65 in 100	14. 50 in 100	50 in 100	or	65 in 100
15. 50 in 100	50 in 100	or	70 in 100	15. 50 in 100	50 in 100	or	70 in 100
16. 50 in 100	50 in 100	or	75 in 100	16. 50 in 100	50 in 100	or	75 in 100
17. 50 in 100	50 in 100	or	80 in 100	17. 50 in 100	50 in 100	or	80 in 100
18. 50 in 100	50 in 100	or	85 in 100	18. 50 in 100	50 in 100	or	85 in 100
19. 50 in 100	50 in 100	or	90 in 100	19. 50 in 100	50 in 100	or	90 in 100
20. 50 in 100	50 in 100	or	95 in 100	20. 50 in 100	50 in 100	or	95 in 100
21. 50 in 100	50 in 100	or	100 in 100	21. 50 in 100	50 in 100	or	100 in 100

neutrality in condition 1.3 (see Andreoni and Sprenger 2011), it does not predict differential risk aversion across conditions 1.3 and 2.3.

Condition 1.3 and 2.3 also help to assess the plausibility that the experimental manipulation alters the structure of the referent, as it is specifically under this assumption that the changing patterns of risk aversion both within condition 1.3 and across conditions 1.3 and 2.3 should arise. The unique prediction of quadratically declining risk aversion in condition 1.3 also allows for a tighter test of the theory than the baseline conditions, as KR will at times predict near risk neutrality and at times predict substantial risk aversion. Further, behavior can be linked across experimental conditions 1.1, 1.2, and 1.3. Hence, one can assess the predictive validity out of sample for KR preference parameters identified under the assumption that the experimental manipulation alters the structure of the referent.²⁸

B. Design Details

In order to eliminate often-discussed confounds (Plott and Zeiler 2005, 2007), neutral language such as “option A” and “option B” was used throughout. The instructions never discussed trading or exchange.

Subjects were told, “In each task you are asked to make a series of decisions between two options: Option A and Option B. In each task Option A will be fixed while Option B will vary. For example, . . . [example]. . . For each row all you have to do is decide whether you prefer Option A or Option B.”

The full instructions are provided as online appendix section A.5. In each condition, the decisions were blocked into tasks corresponding to the three subconditions discussed above. New instructions were read at the beginning of each task block explaining the new procedures and encouraging subjects to make each decision carefully. Subjects were provided with calculators should they wish to use them in making their decisions.

Two orders of the tasks were used in each condition to examine order effects: X.1, X.3, X.2 and X.2, X.3, X.1. The uncertainty equivalents were left in the middle as a buffer between the more similar tasks. No order effects were observed (see online app. table A2). In addition to varying the order, an attempt was also made to manipulate slightly the physical representation of option A in each decision. This was done for around half of the subjects by stapling miniature copies of the appropriate number of bills, or bills with appropriate percentages at the top of each

²⁸ No such exercise can be conducted for condition 2 as KR preference parameters are not identified from certainty equivalent behavior. As noted above, risk neutrality is predicted for all levels of loss aversion in the tasks of condition 2.

decision sheet. The stapling was done such that subjects would be forced to hold the representation of option A in order to make the first few decisions. Though I imagined that this nuance might influence the degree of attachment to option A, it had virtually no effect (see online app. table A2).²⁹ A total of 136 subjects participated in the study across 10 experimental sessions. Table 3 provides the dates, times, orders, and details of all sessions.

In order to provide incentives for truthful revelation of preferences, subjects were randomly paid for one of their choices. The instructions fully described the payment procedure and the mechanism for carrying out randomization of payments, two 10-sided dice. The randomization was described in independent terms. That is, mention was made of rolling dice first for option A and then for option B and an example was given. Subjects earned, on average, \$23 from the study including a \$5 minimum payment that was added to all experimental earnings.

The random-lottery incentive mechanism, which is widely used in experimental economics, does introduce a compound lottery to the decision environment.³⁰ In a series of experiments involving decisions over risky prospects, Starmer and Sugden (1991) and Cubitt, Starmer, and Sugden (1998) demonstrate that this mechanism generally does not suffer from contamination effects in practice as subjects treat decisions effectively in isolation. Online appendix section A.1 and Section IV.A examine the possibility that individuals do not treat each decision in isolation and the corresponding implications for the presence of an endowment effect for risk in both within- and between-subjects data.

IV. Results

The results are presented in three subsections. The first subsection provides a summary of the elicited risk attitudes and tests demonstrating risk aversion in the probability equivalent tasks and virtual risk neutrality in the certainty equivalent tasks. The second subsection is devoted to discussing within-subjects results with data from Andreoni and Sprenger

²⁹ Andreoni and Sprenger (2011) use uncertainty equivalents to test expected utility and investigate violations of first-order stochastic dominance near to certainty. In the non-representation treatments for condition 1.3, the findings are reproduced. However, in the representation treatments for condition 1.3, stochastic dominance violations at certainty are reduced to zero. See Secs. IV.B and IV.C.1 for discussion.

³⁰ This randomization device introduces a compound lottery to the decision environment as individuals each made over 400 choices in their 22 tasks. Reduction of compound lotteries does not change the general equivalence predictions for standard expected utility, prospect theory, and disappointment aversion discussed above. However, to the extent that the compound lottery changes perceived referents, the randomization introduces complications into the KR analysis as it creates a potential link between choices and referents across tasks.

TABLE 3
EXPERIMENTAL SESSIONS

Date	Time	Condition	Order	Representation	Observations
1. May 11, 2010	12:00 p.m.	1	(1) X.1, X.3, X.2	No	10
2. May 11, 2010	2:30 p.m.	1	(2) X.2, X.3, X.1	No	12
3. May 12, 2010	12:00 p.m.	2	(1) X.1, X.3, X.2	No	19
4. May 12, 2010	2:30 p.m.	2	(2) X.2, X.3, X.1	No	16
5. May 18, 2010	12:00 p.m.	1	(1) X.1, X.3, X.2	No	11
6. May 18, 2010	2:30 p.m.	1	(2) X.2, X.3, X.1	No	6
7. May 25, 2010	12:00 p.m.	1	(1) X.1, X.3, X.2	Yes	15
8. May 25, 2010	2:30 p.m.	1	(2) X.2, X.3, X.1	Yes	16
9. May 26, 2010	12:00 p.m.	2	(1) X.1, X.3, X.2	Yes	15
10. May 26, 2010	2:30 p.m.	2	(2) X.2, X.3, X.1	Yes	16
Total					136

NOTE.—“Representation” refers to whether or not option A was physically represented by stapling miniature bills or bills and percentages to the decision sheet.

(2011), which also demonstrate significant differences in elicited risk preferences between certainty and probability equivalent techniques. Third, motivated by these nonparametric results, the KR utility model is estimated and compelling out-of-sample predictions for uncertainty equivalent tasks at both the aggregate and individual levels are provided.

A. Risk Attitudes

Of the 136 individuals who participated in the primary experiment, 70 individuals participated in condition 1 and 66 participated in condition 2. As in most price list style experiments, a number of subjects switch from option A to option B and then back to option A.³¹ Three subjects (4.3 percent) in condition 1 and 11 subjects (16.7 percent) in condition 2 featured multiple switch points in at least one task. The majority of multiple switching occurred in condition 2.3, indicating that this task

³¹ Around 10 percent of subjects feature multiple switch points in similar price list experiments (Holt and Laury 2002; Meier and Sprenger 2010), and as many as 50 percent in some cases (Jacobson and Petrie 2009). Because such multiple switch points are difficult to rationalize and may indicate subject confusion, researchers often exclude such observations or mechanically enforce single switch points. See Harrison et al. (2005) for discussion.

may have been confusing to subjects.³² Attention is given to the 122 subjects who had unique switch points in all 22 decision tasks.³³ This results in 1,474 individual decisions in condition 1 and 1,210 decisions in condition 2. Of these 2,684 total decision tasks, in a small percentage (0.60 percent), the subject preferred option A for all rows and in a larger percentage (4.14 percent) the subject preferred option B for all rows. These responses provide only one-sided bounds on the interval of the subject's response. The other bound is imputed via top and bottom coding accordingly.³⁴

A variety of demographic, cognitive, and attitudinal data were collected after the study was concluded in order to provide a balancing test and be used in control. Table 4 compares data across experimental conditions for survey respondents.³⁵ Though some differences do exist, particularly in academic year, subjects were broadly balanced on observable characteristics, simple numeracy and cognitive ability scores, and subjectively reported risk attitudes.³⁶ An omnibus test from the logit regression of condition assignment on all survey variables for 111 of 122 individuals with complete survey data does not reject the null hypothesis of equal demographic, cognitive, and attitudinal characteristics across conditions ($\chi^2 = 14.1$, $p = .12$). Because the randomization is at the session level, this helps to ensure that accidental selection issues are not driving the experimental results. The results are robust to controlling for demographic differences, and additional within-subjects results are provided in Section IV.B.

³² Five of 11 multiple switchers in condition 2 had multiple switching in only condition 2.3, one had multiple switching in conditions 2.1 and 2.3, two had multiple switching in conditions 2.2 and 2.3, and three had multiple switching in all three subconditions.

³³ All results are maintained when including multiple switchers and taking their first switch point as their choice. See online app. table A2 for details.

³⁴ For example, if an individual chose option A at all rows in a probability equivalent including when option B was a 100 percent chance of getting \$30, I top-code the interval as [100, 100]. Virtually all the bottom-coded responses, 100 of 111, were decisions in condition 2.3 in which the bottom-coded choice would be preferring \$10 with certainty to a given gamble over \$0 and \$30. No bottom-coded responses arose in condition 1 or 2.1, where the lowest option B outcome was \$0 with certainty.

³⁵ Of 122 subjects, 111 completed all survey elements: 60 of 67 subjects in condition 1 and 51 of 55 subjects in condition 2 provided complete survey responses. Nonresponse is unrelated to condition as condition 1 accounts for 54–55 percent of the data in both the respondent and full samples.

³⁶ Numeracy is measured with a six-question exam related to simple math skills such as division and compound interest previously validated in a number of large and representative samples (Banks and Oldfield 2007; Lusardi and Mitchell 2007; Gerardi, Goette, and Meier 2010). Cognitive ability is measured with the three-question Cognitive Reflection Test introduced and validated in Frederick (2005). Subjective risk attitudes are measured on a 7-point scale with the question "How willing are you to take risks in general on a scale from 1 (unwilling) to 7 (fully prepared)" previously validated in a large representative sample (Dohmen et al. 2005).

TABLE 4
SUMMARY STATISTICS AND BALANCING TEST

VARIABLE	OBSERVATIONS	MEAN (Standard Deviation)			<i>t</i> - STATISTIC	<i>p</i> - VALUE
		Total (<i>N</i> = 122)	Condition 1	Condition 2		
			(<i>N</i> = 67)	(<i>N</i> = 55)		
Male (= 1)	119	.46 (.50)	.42 (.50)	.52 (.50)	1.12	.27
Academic year	122	2.63 (1.08)	2.45 (1.05)	2.85 (1.08)	2.10	.04
Grade point average	120	3.20 (.42)	3.25 (.44)	3.15 (.39)	−1.26	.21
English 1st language (= 1)	122	.56 (.50)	.51 (.50)	.62 (.49)	1.22	.22
Smoker (= 1)	122	.04 (.20)	.03 (.17)	.05 (.23)	.68	.50
Weekly spending (\$)	122	89.68 (89.49)	85.00 (68.63)	95.38 (110.12)	.64	.53
Risk attitudes (1–7)	122	3.84 (1.19)	3.70 (1.22)	4.00 (1.14)	1.39	.17
Cognitive ability score (1–3)	117	1.79 (1.05)	1.86 (1.08)	1.70 (1.01)	−.83	.41
Numeracy score (1–6)	120	5.75 (.54)	5.78 (.52)	5.71 (.57)	−.76	.45

NOTE.—Omnibus $\chi^2 = 14.1$ ($p = .12$). The table presents summary statistics for 122 subjects with unique switch points in all 22 decision tasks. Observations refers to the number of responses to each question. Omnibus χ^2 test statistic corresponds to the null hypothesis of zero slopes in logit regression with 111 subjects with complete survey data of condition assignment on all survey variables with robust standard errors.

I begin by investigating behavior in conditions 1.1, 1.2, 2.1, and 2.2. With the exception of the KR preference model, all discussed theories predict experimental equivalence across these conditions. That is, elicited risk attitudes should be identical whether one asks the probability equivalent of a given certain amount (conditions 1.1 and 1.2) or the certainty equivalent of a given gamble (conditions 2.1 and 2.2). To graphically present the data, I first define the risk premium,

$$\text{Risk Premium} \equiv \begin{cases} E(PE^*) - C & \text{if condition} = 1 \\ EV - CE^* & \text{if condition} = 2, \end{cases}$$

where the starred terms represent the midpoint of the interval implied by a subject's switch point in either the probability equivalents (PE) of condition 1 or the certainty equivalents (CE) of condition 2. The term C refers

to the experimentally given certain amount in condition 1, $E(\cdot)$ is the expectation operator, and EV refers to the expected value of the experimentally given gamble in condition 2. A positive risk premium implies risk aversion. Figure 2 presents the risk premia for the 122 individuals with unique switch points. Median data, 25th–75th percentile interquartile ranges, and 5th–95th percentile bars are provided. The experimental parameters, C and EV , are presented on the horizontal axis and risk premia are presented on the vertical axis.

Apparent from the median data is the systematic difference in elicited risk attitudes between certainty and probability equivalents. When fixing a binary gamble and choosing over changing certain amounts in conditions 2.1 and 2.2, subject median data display virtual risk neutrality with risk premia near zero. When fixing a certain amount and choosing over changing gambles in conditions 1.1 and 1.2, subjects display risk aversion with substantial positive risk premia. Further, the interquartile ranges suggest that a large majority of individuals exhibit risk aversion for each probability equivalent question but that risk aversion is more muted and potentially related to the assessment probability for each certainty equivalent.³⁷

The messages obtained from graphical differences in risk premia across conditions are echoed in more coarse analyses as well. For each experimental task, the corresponding decision can be classified as being risk neutral, risk averse, or risk loving. These classifications recognize the interval nature of the data. For example, a decision is coded as risk neutral if the risk-neutral response lies in the interval generated by the subject's switch point. Table 5 provides a tabulation of these classifications. Proportionally, nearly twice as many observations are classified as risk averse in the probability equivalents of condition 1 compared to the certainty equivalents of condition 2.³⁸

The raw data provide compelling evidence that both the direction and level of risk attitudes differ when comparing certainty equivalents and probability equivalents, consistent with the KR endowment effect for risk. To understand both the magnitude and significance of the results,

³⁷ The observed certainty equivalent pattern is consistent with cumulative prospect theory's S-shaped probability weighting function (Kahneman and Tversky 1979; Tversky and Kahneman 1992). In cumulative prospect theory, low probabilities are up-weighted and high probabilities are down-weighted such that certainty equivalents can exhibit risk loving at low probabilities and risk aversion at high probabilities. In experiments eliciting certainty equivalents for gambles, Tversky and Kahneman (1992) and Tversky and Fox (1995) estimated utility parameters supporting such a view.

³⁸ As this may be a strict classification of responses, one can also extend the interval of the switch point to plus or minus one choice, drawing classification toward risk neutrality. By this wider interval measure, in condition 1, 56 percent of observations are classified as risk averse vs. 28 percent for condition 2. Additionally, 54 percent of the condition 2 data are classified as risk neutral.

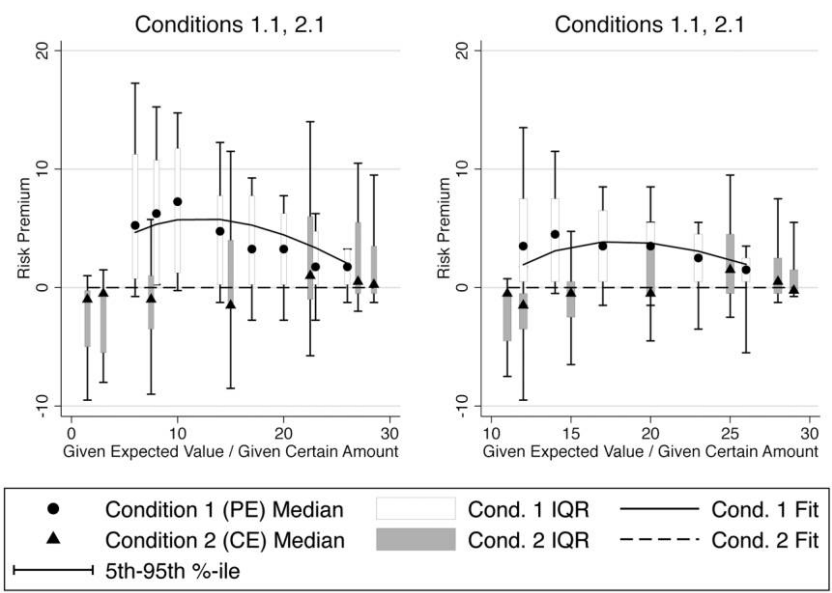


FIG. 2.—Conditions 1.1, 1.2, 2.1, and 2.2 responses. Data are from 122 experimental subjects with unique switching points in all 22 decision tasks. Dots correspond to medians, rectangles to interquartile ranges (IQR), and bars to the 5th–95th percentiles. Lines correspond to KR models fit with $\lambda = 3.4$. The KR model predicts risk aversion for probability equivalents in conditions 1.1 and 1.2 and risk neutrality for certainty equivalents in conditions 2.1 and 2.2.

as well as control for potential confounds, table 6 provides two sets of regressions. Columns 1–3 present interval regressions (Stewart 1983) with the dependent variable as the interval of risk premium implied by a subject’s switch point. Columns 4–6 present ordered logit regressions. The dependent variable is risk attitude, which takes the value -1 for a

TABLE 5
CLASSIFICATION OF RISK ATTITUDES

	Condition 1 Probability Equivalents	Condition 2 Certainty Equivalents
Risk averse	628 (.67)	292 (.38)
Risk neutral	222 (.24)	263 (.34)
Risk loving	88 (.09)	215 (.28)
Total	938 (1.00)	770 (1.00)

NOTE.—The table presents the classification of risk attitude for each decision based on interval of subject switch points. Entries are the number (proportion) of observations reported for 122 subjects without multiple switching over 22 experimental tasks.

risk-loving classification, 0 for risk neutrality, and +1 for a risk-averse response. The natural order of risk attitude corresponds to increasing risk aversion. These regressions control for both condition assignment and the dollar value of the experimentally varied parameter, C in the case of probability equivalents or EV in the case of certainty equivalents.³⁹ Hence, the regressions examine, for a common stake size, the degree to which risk preferences are influenced by elicitation technique.

The results indicate a substantial effect of condition assignment. Controlling for the dollar value of the experimentally varied parameter, assignment to probability equivalents induces \$3.50 greater risk premia relative to certainty equivalents. Interestingly, assignment to certainty equivalents generates a risk premium that is, on average, indistinguishable from zero. Column 4 echoes this message. Individuals assigned to probability equivalents are significantly more likely to exhibit risk aversion. Odds ratios for being classified as risk averse relative to risk neutral or risk loving are provided in brackets. Subjects randomly assigned to condition 1 are around 3.5 times more likely to exhibit risk aversion than those assigned to condition 2. Importantly, the specifications of columns 2–3 and 5–6 additionally control for order and representation effects as well as the collected demographic and attitudinal characteristics for individuals who responded in full to the poststudy survey. Across specifications, subjects in condition 1 are significantly more likely to exhibit risk aversion.⁴⁰

These simple tests are suggestive of an endowment effect for risk. In certainty equivalents tasks, subjects are generally risk neutral. In probability equivalents tasks, subjects are generally risk averse. Standard theories and DA predict experimental equivalence across these two environments. The data are potentially consistent with the KR model, with its possibility of a stochastic reference distribution. If the fixed decision element is perceived as the referent, the KR model predicts precisely the observed pattern of behavior. Another possibility is that such an effect is generated as a KR PPE outcome in an environment in which subjects face different choice sets across experimental conditions. Next, I consider within-subject data in which each individual faces the same set of choices, which may help to distinguish between these two possibilities.

B. Within-Subjects Data

In this experiment it may be natural to envision decision making being organized around the experiment as a whole. If the rational expecta-

³⁹ For reference, the mean value of C is \$17.50 while the mean value of EV is \$16.86.

⁴⁰ These results are maintained with the inclusion of multiple switchers. Additionally, no interactions for order or representation effects were obtained. Online app. table A2 provides these additional regressions.

TABLE 6
PROBABILITY AND CERTAINTY EQUIVALENT RISK PREFERENCE REGRESSIONS

	INTERVAL REGRESSION: DEPENDENT VARIABLE RISK PREMIUM INTERVAL			ORDERED LOGIT: DEPENDENT VARIABLE RISK ATTITUDE CLASSIFICATION		
	(1)	(2)	(3)	(4)	(5)	(6)
PE: condition 1 (= 1)	3.588 ^{***} (.473)	3.568 ^{***} (.470)	3.355 ^{***} (.489)	1.269 ^{***} (.219) [3.557]	1.268 ^{***} (.218) [3.554]	1.110 ^{***} (.235) [3.034]
C _{PE} or EV _{CE}	.014 (.026)	.014 (.026)	.006 (.027)	.031 ^{***} (.009)	.031 ^{***} (.009)	.029 ^{**} (.010)
Male (= 1)			-.119 (.524)			-.086 (.237)
Academic year			-.212 (.225)			-.223 [*] (.094)
Grade point average			-.056 (.602)			.275 (.267)
English 1st language (= 1)			-1.084 [*] (.528)			-.368 (.249)
Smoker (= 1)			-.160 (1.509)			-.322 (.505)
Weekly spending (\$)			.002 (.002)			.001 (.001)

tions referent is constructed as the distribution of outcomes induced by all experimental choices, KR PPE could be a viable organization of the between-subjects data. However, such a construction cannot deliver a difference between elicited risk aversion across probability and certainty equivalents in a within-subjects design. See online appendix A.1 for further detail. In this section, within-subjects data are examined, which may be helpful for distinguishing between the more broadly framed PPE logic and the nonequilibrium rationale that the fixed decision element is perceived as the referent. More generally, these data may also be used to analyze the sensitivity of the observed results to the altered protocol of a between- versus within-subjects design.

I present portions of data obtained for Andreoni and Sprenger (2011), a within-subjects study of 76 subjects designed primarily with uncertainty equivalents similar to condition 1.3 and certainty equivalents similar to condition 2.1.⁴¹ Importantly, in one task subjects are asked to state the gamble probability over \$30 and \$0 that makes them indifferent to a 100 percent chance of receiving \$10. This probability equivalent can be compared to the subject's own certainty equivalents by finding the gamble over \$30 and \$0 at which the certainty equivalent is revealed to be \$10.⁴²

In certainty equivalents, the average gamble over \$30 and \$0 at which the certainty equivalent is revealed to be \$10 is 0.306 (standard deviation 0.210), implying a risk premium of $-\$0.82$ (6.31). In contrast, the average (q ; 30, 0) probability equivalent of \$10 for sure is 0.541 (0.220), implying a risk premium of $\$6.22$ (6.61). The difference in risk premia is significant at all conventional levels ($t = 6.72$, $p < .001$).⁴³ Within-

⁴¹ Some minor experimental differences exist such as the use of prechecked boxes as decision aids, differing orders, and no representation in Andreoni and Sprenger (2011). Two Holt and Laury (2002) risk tasks were implemented as a buffer between uncertainty equivalents and certainty equivalents in Andreoni and Sprenger (2011) and decisions were collected between conditions. See Sec. IV.C.1 and Andreoni and Sprenger's study for more details.

⁴² For certainty equivalents tasks, the gamble that yields a certainty equivalent of \$10 is identified by finding the smallest gamble probability, p , at which the certainty equivalent is higher than \$10, finding the largest probability, p' , at which the certainty equivalent is smaller than \$10, and taking the average $\bar{p} = (p + p')/2$. There are 13 cases in which p' is not observed because the individuals' certainty equivalents always exceeded \$10. There is one case in which p is not observed because the individual's certainty equivalent never exceeded \$10. In these cases the nonmissing value is used. Eliminating these observations does not influence the result.

⁴³ These values are remarkably consistent with the between-subjects evidence documented in Secs. IV.A and IV.C.1, where subjects are close to risk neutral in certainty equivalents and reveal average probability equivalents for \$10 with certainty of 0.547 (0.189) in condition 1.3 and 0.546 (0.185) in condition 1.1. Interestingly, the KR model with the standard values of $\lambda = 3$, $\eta = 1$ and the assumption that the referent is perceived as the fixed element of a task would predict a risk premium of \$0 in the implemented certainty equivalents and a risk premium of \$5 in the implemented probability equivalent. The within-subjects data are therefore consistent with the KR preference model with standard values and those obtained in the between-subjects estimates of Sec. IV.C.1.

subjects probability equivalents generate substantially more risk aversion than certainty equivalents. Fixing a choice between \$10 and a 50–50 gamble over \$0 and \$30, in a probability equivalent, 54 percent of subjects prefer the \$10, while in a certainty equivalent, 68 percent prefer the gamble.⁴⁴

In the within-subject design, the same subjects are close to risk neutral in certainty equivalents and significantly more risk averse in probability equivalents. This reproduction of the central between-subjects finding helps to ensure that the protocol of a between-subjects design with different subjects facing different choice sets is not driving the results. Importantly, these results help to distinguish between nonequilibrium first-focus behavior and the PPE refinement with a referent constructed as the distribution of outcomes induced by all experimental choices. In a within-subjects experiment, the KR model in PPE cannot generate the inconsistencies required of the endowment effect for risk.⁴⁵ Hence, these results provide helpful distinction, ruling out PPE and maintaining first focus as a valid organization for the data.

C. Model Estimation and Predictive Validity

Under the first-focus assumption that the referent is the fixed element in each decision environment, the KR model rationalizes the data. When choosing between a fixed certain amount and changing gambles, the fixed certain element acts as the referent, delivering risk aversion. When choosing between a fixed gamble and changing certain amounts, the fixed gamble acts as the referent, delivering risk neutrality.

I now turn to estimating the KR model using first focus as an identifying assumption. In the following subsections, I estimate the KR model and assess the predictive validity of the resulting parameters both in and out of sample at the aggregate and individual levels. These exercises help to evaluate the plausibility of the notion that the experimental manipulations do indeed alter the stochastic structure of the referent.

1. Estimating Aggregate KR Preferences

Under the first-focus identifying assumption that the referent is the fixed element in each decision environment, the KR model is easily

⁴⁴ This is a conservative estimate of the preference for the gamble in the certainty equivalent as the experimental choice itself was between \$10.50 and a 50–50 gamble over \$0 and \$30 and 68 percent of subjects chose the gamble.

⁴⁵ An alternative between-subjects experimental design would be to have subjects face the exact same choices but organized differently into a probability equivalent or a certainty equivalent organization. As subjects face the same choices, PPE logic with a referent taken as the distribution of outcomes induced by choice could not rationalize the endowment effect for risk.

implemented econometrically. In this subsection, I return to the primary between-subjects data and attempt to both estimate such a model and assess its predictive validity in sample.

With the identifying assumption that the referent, τ , is the fixed decision element and experimental variation in x_1 and x_2 , structural estimation is straightforward. Assuming an additive error, equation (3) can be estimated using standard techniques with λ recovered as the parameter of interest.⁴⁶ Using the data from probability equivalent conditions 1.1 and 1.2, the midpoint of the interval implied by a subject's switch point is taken as the value q^* in equation (3), which is then estimated via nonlinear least squares with standard errors clustered on the subject level.⁴⁷ The aggregate estimate is $\hat{\lambda} = 3.41$ (standard error [SE] = 0.34). The null hypothesis of zero loss aversion, $\lambda = 1$, is rejected $F_{1,66} = 49.06$, $p < .01$. This value of loss aversion is consistent with loss aversion estimates from other contexts (Tversky and Kahneman 1992; Pope and Schweitzer 2011; Gill and Prowse 2012) and is closely in line with the often-discussed benchmark of losses being felt twice as severely as gains, $\lambda = 3$, $\eta = 1$ (Koszegi and Rabin 2006, 2007).⁴⁸ The in-sample *R*-squared value for the estimation sample of condition 1 is .94 and the correlation between predicted and actual probability equivalents is .74. Figure 2 presents predicted values for risk premia from this aggregate regression as the solid lines for conditions 1 and 2.⁴⁹ The aggregate data match well the model's predictions in both conditions. Precisely matching the aggregate data patterns demonstrates the in-sample predictive validity of the KR preference model under the assumption that the fixed decision element is perceived as the referent. In the next subsection we analyze predictive validity out of sample in conditions 1.3 and 2.3.

⁴⁶ That is, the estimating equation is

$$q^* = \frac{r - x_2 - \lambda \cdot (x_2 - r)}{(x_1 - x_2) + [1 \cdot (x_1 - r) - \lambda \cdot (x_2 - r)]} + \epsilon.$$

⁴⁷ As no measure of loss aversion is recoverable from condition 2 under the first-focus assumption, this portion of the data is ignored when estimating preferences at the aggregate and individual levels. Hence, there is no opportunity to examine heterogeneous treatment effects over types as in, e.g., Goette and Fehr (2007).

⁴⁸ The functional form of Tversky and Kahneman (1992) does not feature consumption utility, and so the loss aversion estimate of $\hat{\lambda} = 2.25$ in their paper is a direct measure of losses being felt twice as severely as gains.

⁴⁹ For condition 1 this is calculated as the implied risk premium from the predicted probability equivalents of the nonlinear least squares regression. For condition 2, the risk premium is zero everywhere as the KR model under the assumption that the fixed decision element is perceived as the referent predicts risk neutrality.

2. Aggregate Out-of-Sample Predictive Validity

In order to evaluate the out-of-sample predictive validity of the KR preference model under the first-focus referent, I analyze behavior in conditions 1.3 and 2.3. As noted above, if the fixed decision element is perceived as the referent, the KR preference model predicts risk neutrality in condition 2.3 and predicts a particular shape of quadratically declining risk aversion in condition 1.3. Figure 3 presents data from these conditions. For each response calculate the expected risk premium as

$$\begin{aligned} &\text{Expected Risk Premium} \\ \equiv &\begin{cases} E(UE^*) - EV & \text{if condition} = 1 \\ EV - E(\text{inverted } UE^*) & \text{if condition} = 2. \end{cases} \end{aligned}$$

The starred terms represent the midpoint of the interval implied by a subject's switch point in either the uncertainty equivalents (UE) of condition 1 or the inverted uncertainty equivalents (inverted UE) of condition 2, $E(\cdot)$ is the expectation operator, and EV refers to the expected value of the experimentally given gamble. The expected risk premium is the direct analogue of the standard risk premium in choices between gambles. Positive expected risk premium implies risk aversion.

Figure 3 highlights a distinction in risk aversion between conditions 1.3 and 2.3, supporting the KR formulation. At lower values of EV , $EV \leq 18$, expected risk premia in condition 1.3 are significantly larger than

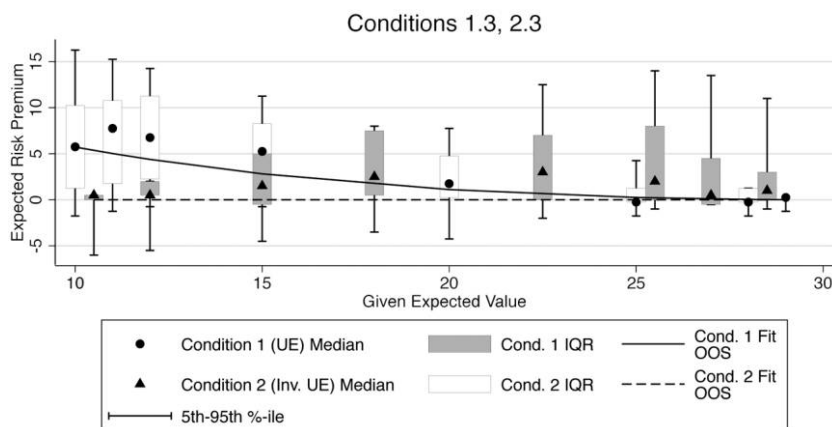


FIG. 3.—Conditions 1.3 and 2.3. Data are from 122 experimental subjects with unique switching points in all 22 decision tasks. Dots correspond to medians, rectangles to inter-quartile ranges (IQR), and bars to the 5th–95th percentiles. Solid lines correspond to out-of-sample KR model prediction with $\lambda = 3.4$ as estimated in conditions 1.1 and 1.2. The KR model predicts risk aversion in condition 1.3 and risk neutrality in condition 2.3.

those of condition 2.3.⁵⁰ However, at higher values of EV , $EV > 18$, risk aversion is muted in both conditions with expected risk premia in condition 1.3 actually being significantly lower than those of condition 2.3.⁵¹

Figure 3 additionally presents out-of-sample predictions for the KR model with the estimated $\hat{\lambda} = 3.4$. Though the KR prediction of risk neutrality breaks down somewhat at the intermediate probabilities of condition 2.3, in condition 1.3, the out-of-sample prediction closely matches aggregate behavior.⁵² The out-of-sample correlation between predicted and actual expected risk premia is .45.⁵³

Matching the aggregate data patterns in conditions 1.3 and 2.3 demonstrates out-of-sample predictive validity for the KR preference model under the first-focus assumption. Indeed, condition 1.3 puts the theory to a stricter test than the baseline conditions. If changing decision elements alter the stochastic structure of the referent, the theory predicts a specific shape of quadratically declining risk premia, at times predicting near risk neutrality and at times risk aversion. In the aggregate this is

⁵⁰ For $EV \leq 18$ in condition 1.3, expected risk premia are 6.33 (clustered SE = 0.58) while for condition 2.3 risk premia are 1.15 (0.32). The difference between the two is significant at all conventional levels, $\chi^2(1) = 61.55$, $p < .01$. Values and test statistic are from interval regression of expected risk premium on condition assignment and a constant for conditions 1.3 and 2.3 with $EV \leq 18$ with standard errors clustered at the individual level.

⁵¹ For $EV > 18$ in condition 1.3, expected risk premia are 0.56 (clustered SE = 0.27) while for condition 2.3 expected risk premia are 3.35 (0.55). The difference between the two is significant at all conventional levels, $\chi^2(1) = 20.76$, $p < .01$. Values and test statistic are from interval regression of expected risk premium on condition assignment and a constant for conditions 1.3 and 2.3 with $EV > 18$ with standard errors clustered at the individual level.

⁵² Additionally, condition 1.3 reproduces the general shape and level of the uncertainty equivalents discussed in Andreoni and Sprenger (2011), demonstrating a slightly convex relationship between given gambles and their uncertainty equivalents. However, Andreoni and Sprenger document the convexity becoming sharper as the given gamble approaches certainty, and this result is not present in the data. Minor differences in experimental detail may account for the difference at $p = 0$ between the present results and those of Andreoni and Sprenger, including a different number and order of tasks and slightly changed tasks. The Andreoni and Sprenger price lists were designed with decision aids of checked top and bottom rows. The task used in condition 1.3 was not. More important, however, appears to be the presence of the physical representation of option A. The sharpened convexity at $p = 0$ in Andreoni and Sprenger's study is driven by individuals who violate first-order stochastic dominance close to certainty. They document individual dominance violation rates between $p = 0$ and $p = .05$ of around 17.5 percent across three tasks. When option A is not physically represented, a similar violation rate of 13.5 percent is found. However, when option A is physically represented, zero violations of stochastic dominance at certainty are observed. The effect of physical representation on the proportion of individuals violating stochastic dominance at certainty is significant ($z = 2.09$, $p < .05$).

⁵³ For condition 1.3 alone, the out-of-sample correlation between predicted and actual expected risk premia is .57. For condition 2.3 alone, the out-of-sample correlation is not calculated as the predicted expected risk premium is always zero. For the underlying uncertainty equivalent responses, one can also calculate out-of-sample correlations. For condition 1.3, the out-of-sample correlation between predicted and actual uncertainty equivalents is .73, while for condition 2.3, the out-of-sample correlation between predicted and actual uncertainty equivalents is .81.

broadly observed. Next, I provide individual estimates of loss aversion for subjects in condition 1 and investigate this prediction in closer detail, with a comparison to plausible alternatives.

3. Individual Out-of-Sample Predictive Validity

For the 67 subjects in condition 1, individual analyses can be conducted. Using the data from conditions 1.1 and 1.2, the degree of loss aversion, λ_i , can be estimated following equation (2).⁵⁴ These individual estimates, $\hat{\lambda}_i$, can be correlated with behavior in condition 1.3. Figure 4, panel A, provides a histogram of the individual estimates of $\hat{\lambda}_i$. The median $\hat{\lambda}_i$ estimated from conditions 1.1 and 1.2 is 3.6, echoing the aggregate result.⁵⁵ Figure 4, panel B, provides a histogram of expected risk premia in condition 1.3 for these same 67 subjects. Each subject has eight such risk premia corresponding to the eight tasks in condition 1.3. The median expected risk premium is 1.25 indicating risk aversion, and the data indicate wide variation in behavior.

Figure 4, panel C, investigates the link between KR preference parameter estimates in conditions 1.1 and 1.2 and the expected risk premia in condition 1.3. For each individual, the estimate of $\hat{\lambda}_i$ is used to predict each response in condition 1.3 and the corresponding expected risk premium. The predicted and actual risk premia are graphed together to examine the extent of predictive validity at the individual level in an environment in which the theory at times predicts near risk neutrality and at times substantial risk aversion. Notable from figure 4, panel C, is the extent to which the actual and predicted expected risk premia accord throughout the domain. The correlation between predicted and actual expected risk premia is .76. The corresponding regression line is close to the 45-degree line with a slope of 1.06 (SE = 0.07) and an intercept of 0.90 (0.32). The null hypothesis that the slope is 1 is not rejected, $F(1, 66) = 0.86$, $p = .36$.⁵⁶ At the individual level, the KR preference model under the first-focus referent adheres closely to behavior out of sample both when the model predicts near risk neutrality and

⁵⁴ For each individual, i , the estimating equation is

$$q^* = \frac{r - x_2 - \lambda_i \cdot (x_2 - r)}{(x_1 - x_2) + [1 \cdot (x_1 - r) - \lambda_i \cdot (x_2 - r)]} + \epsilon,$$

which is estimated via nonlinear least squares.

⁵⁵ The individual estimates provide increased in-sample predictive validity relative to the aggregate estimator. The correlation between predicted and actual probability equivalents of conditions 1.1 and 1.2 based on the individual-level predictions is .91. Recall that the in-sample correlation between predicted and actual probability equivalents based on the aggregate predictions was .74.

⁵⁶ However, the null hypothesis that the intercept is 0 is rejected, $F_{1,66} = 8.14$, $p < .01$. Estimates and hypothesis tests are from regression of the predicted premia on actual premia with coefficients clustered at the individual level.

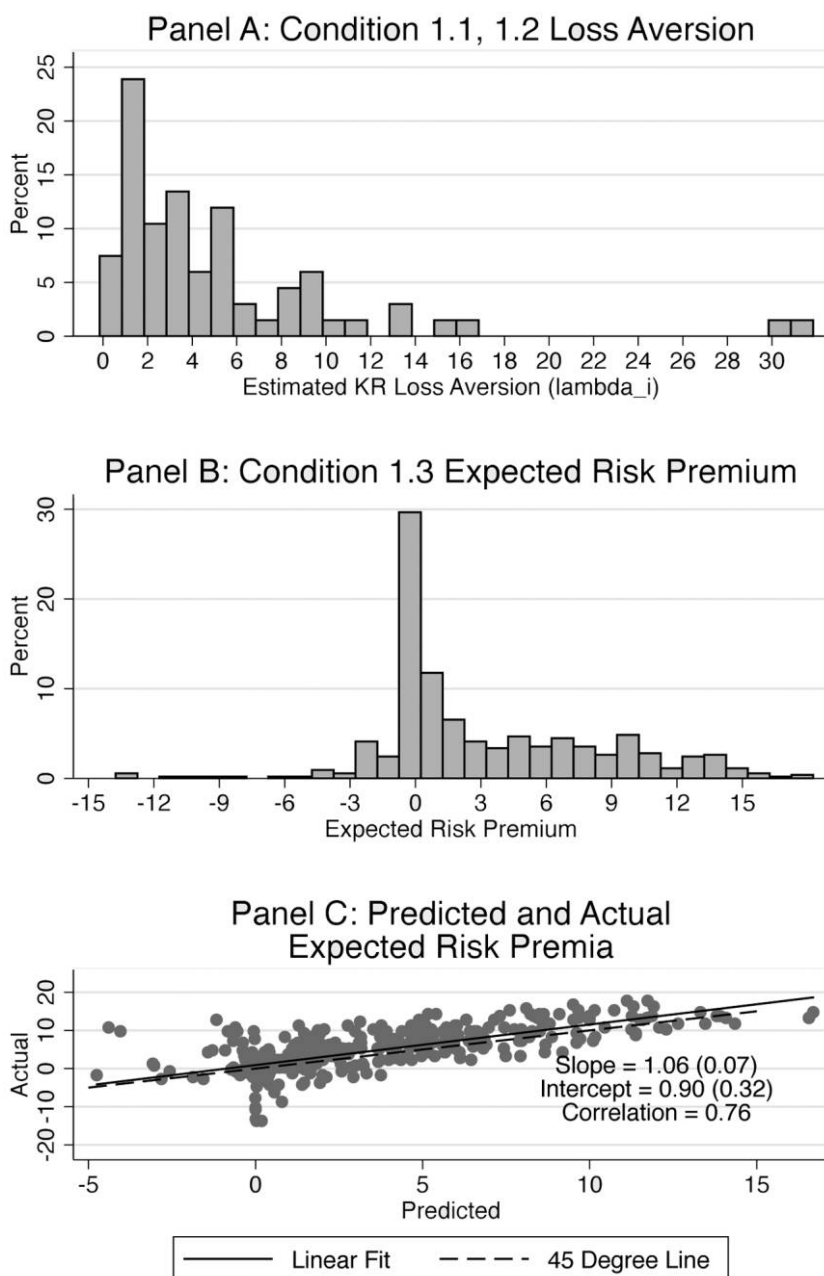


FIG. 4.—Individual loss aversion ($\hat{\lambda}_i$) and expected risk premium. Individual loss aversion estimate is from $\hat{\lambda}_i$ of equation (2) using data from conditions 1.1 and 1.2. Expected risk premia are calculated from condition 1.3 responses. Predicted expected risk premia for condition 1.3 are calculated for each individual on the basis of estimated $\hat{\lambda}_i$.

when the model predicts risk aversion. In online appendix figure A2, figure 4, panel C, is separated for each task in condition 1.3 demonstrating that the actual and predicted risk premia correspond closely across tasks.

Demonstrating the predictive validity of the implemented KR model under the first-focus referent both in and out of sample at the aggregate and individual levels when the model predicts risk aversion and when the model predicts risk neutrality helps support the plausibility that the experimental manipulation does indeed alter the structure of the referent. Naturally, one may wish to consider alternative formulations for comparison.

A first question is how would a standard expected utility (EU) formulation perform in a similar suite of comparisons? Clearly, an expected utility formulation will be unable to rationalize the central treatment effect of differing risk aversion across probability and certainty equivalents. However, for conditions 1.1 and 1.2, one can still conduct the exercise, estimating or calculating a level of EU risk aversion and predicting out of sample for condition 1.3. I conduct two such exercises. First, I assume constant relative risk-averse utility, $u_i(x) = x^{\alpha_i}$, and estimate α_i for each individual on the basis of conditions 1.1 and 1.2.⁵⁷ These estimates are then used to predict out of sample to condition 1.3. The correlation between predicted and actual expected risk premia is .75, slightly lower than the first-focus correlation. Online appendix figure A3, panel A, plots the data. The corresponding regression line has a slope of 0.87 (SE = 0.07) and an intercept of 0.38 (0.41). The null hypothesis that the slope is 1 is marginally rejected, $F_{1,66} = 3.43$, $p = .07$. Second, instead of positing a functional form for utility, I calculate the cardinal utility index for receiving \$10, $u(10)$, directly from the probability equivalent of \$10 in condition 1.1. Under the EU assumptions one can normalize, $u(30) = 1$ and $u(0) = 0$, such that this probability equivalent is exactly the cardinal index. On the basis of this calculation one can predict all responses in condition 1.3 as the uncertainty equivalents involve only \$0, \$10, and \$30 as potential outcomes. Using this method, the correlation between predicted and actual expected risk premia is .74, again lower than the first-focus correlation. Online appendix figure A3, panel B, plots the data. The corresponding regression line has a slope of 0.85 (SE = 0.07) and an intercept of 0.39 (0.39). The null hypothesis that the slope is 1 is rejected, $F_{1,66} = 5.01$, $p = .03$. There is a simple

⁵⁷ For each individual, i , the estimating equation is

$$q^* = \frac{r^{\alpha_i} - x_2^{\alpha_i}}{x_1^{\alpha_i} - x_0^{\alpha_i}} + \epsilon,$$

which is estimated via nonlinear least squares. The variable r here refers to the fixed certain amount in the probability equivalent even though there is no explicit referent.

explanation for why both EU-based exercises generate predicted-actual slope coefficients lower than one. Because risk aversion is consistent under EU, the model cannot match the true declining pattern of risk aversion, predicting more risk aversion than truly exists when expected risk premia are low and less than truly exists when expected risk premia are high.

A second question is, how would a less structured behavioral model perform out of sample? One such model of behavior is that individuals are drawn to the fixed option A in each task because of salience or for other heuristic reasons. Such effects might lead subjects to specific decision rules, perhaps inducing subjects to switch a fixed number of rows below expected value in each task. Given the organization of the tasks, such a salience-based rule would lead to risk aversion in probability equivalents and risk tolerance in certainty equivalents.⁵⁸ This generates an asymmetry in elicited risk aversion, though not precisely the observed endowment effect for risk that entails risk neutrality in certainty equivalents. For conditions 1.1 and 1.2, I estimate such a fixed-row decision rule for all individuals and predict out of sample to condition 1.3 under the assumption that they maintain the same rule. The correlation between predicted and actual expected risk premia is .65, substantially lower than the first-focus correlation. The reason for this diminished correlation is that if individuals implement a fixed-row decision rule, their expected risk premia will be constant across the tasks of condition 1.3. Online appendix figure A3, panel C, plots the data. Interestingly, though such a model misses all the individual-level variation in behavior, the corresponding regression line has a slope of 1.02 (SE = 0.10) and an intercept of 0.02 (0.49). The null hypothesis that the slope is 1 is not rejected, $F_{1,66} = 0.03$, $p = .86$.

The conclusion of these exercises is that the KR model, under the assumed first-focus referent, outperforms standard EU formulations and unstructured decision rule formulations in predictive validity in terms of correlation and quality of point predictions. This is encouraging evidence supporting the first-focus rationale and the interpretation that experimental manipulation does indeed alter the stochastic structure of the referent.

V. Discussion and Conclusion

The obtained results are supportive of the KR preference model. Both between and within subjects the data indicate an endowment effect for risk. In the primary study, subjects are between three and four times

⁵⁸ The reason is that holding on to option A in a certainty equivalent for a fixed number of rows beyond expected value leads to risk tolerance while the same rule in a probability equivalent leads to risk aversion. See table 1 for detail.

more likely to be risk averse if randomly faced with a probability equivalent as opposed to a certainty equivalent. In a secondary, within-subjects study, the phenomenon is also observed. Under the assumption that the experimental variation changes the perceived referent, the KR preference model organizes the data at both the aggregate and individual levels and both between and within subjects, in and out of sample.

Unlike prior work demonstrating the importance of expectations for reference dependence, these results are able to distinguish between KR preferences and other expectations-based models such as disappointment aversion. Gaining separation between these models is necessary for evaluating theoretical developments that depend critically on the stochasticity of the referent (Koszegi and Rabin 2006, 2007, 2009; Heidhues and Koszegi 2008; Herweg, Muller, and Weinschenk 2010).

Additionally, the distinction between the KR and DA models is important in applied settings in which the two models make different predictions. A brief set of simple applications in finance, insurance purchasing, and decision science are considered. Caution should be taken in considering these applications as the links between these experimental findings with manipulated referents and real-world applications in which the referent should exist naturally are not established.

These applications build from the feature of behavior that most prominently distinguishes KR from DA: that in KR the amount of risk a decision maker is willing to take depends on the riskiness of the referent. Consequently, though the DA model predicts first-order risk aversion in the sense of Segal and Spivak (1990) over all binary gambles, the KR model predicts first-order risk aversion for binary gambles only when the referent is relatively certain.⁵⁹

First, in finance, first-order risk aversion is argued to influence stock market participation (Haliassos and Bertaut 1995; Barberis, Huang, and Thaler 2006) and returns (Epstein and Zin 1990; Barberis and Huang 2001). The KR model indicates that individuals may be more likely to accept fair bets that lie within referent expectations but grow more risk averse if outcomes fall outside the referent. As such, the KR model would predict increased risk aversion, with corresponding participation and returns effects when potential outcomes lie outside the referent. In contrast, in DA, individuals are first-order risk averse over all gambles, never accepting a fair investment bet. Second, a comparison can be made for insurance purchasing, where first-order risk aversion potentially influences contract choice (Sydnor 2010). First-order risk aversion predicts a

⁵⁹ Near risk neutrality is also maintained in the KR model when the referent is binary but gamble outcomes are more variable. In the KR model, a relationship between levels of risk aversion and expectations extends beyond binary environments but can no longer be expressed in terms of orders of risk aversion as decision makers will always evaluate over first-order kink points.

desire for full insurance even when insurance is not actuarially fair (Segal and Spivak 1990). Hence, a disappointment-averse consumer may fully insure under positive insurance profits. The KR model predicts that insurance will be more likely for potential outcomes outside of referent expectations. As such, the KR model gives a suggestion as to what gambles will be insured and how they relate to the referent.⁶⁰

Third, a gap between KR and DA models is apparent in decision science, where researchers have long debated the inconsistency between probability equivalent and certainty equivalent methods for utility assessment (Hershey et al. 1982; Hershey and Schoemaker 1985; McCord and de Neufville 1985, 1986; Schoemaker 1990). The general finding is that probability equivalents yield more risk aversion. This difference in elicited utility is predicted by the KR model but not by DA, provided that the utility assessment method alters the referent. Interestingly, the literature indicated that such inconsistency may be due to a “response mode bias,” where probability equivalent tasks are reframed as mixed gambles with gains and losses (Hershey et al. 1982; Hershey and Schoemaker 1985), though other explanations for the inconsistency were proposed (McCord and de Neufville 1985; Schoemaker and Hershey 1992; Schoemaker 1993). The present results and use of the KR model may help to resolve this long-standing issue. The specific KR formulation of preferences and not some idiosyncratic bias organizes this long-debated inconsistency in utility elicitation.

Though the findings are supportive of the KR model and provide direction for further analysis, a distinction must be made between the results presented and the equilibrium predictions of the KR model. The data usefully distinguish between the KR rational expectations equilibrium refinements and alternative nonequilibrium rationales. The data are sensibly organized by the nonequilibrium “first-focus” referent proposed by Koszegi and Rabin (2006) both in and out of sample.

The observed sensitivity of referents to contextual changes has implications for both economic agents and experimental methodology. First, from a methodological perspective, if fixed elements can serve as referents, slightly changed choice environments may induce very different behavior. This is of particular importance for the experimental measurement of preferences such as risk and time preferences, where price list techniques are often used and resulting estimates are given economic significance. The findings suggest that inconsistencies will

⁶⁰ Of course, risks can lie outside of referent expectations by endogenously constructing a fixed referent, and risks can coincide with expectations by constructing a stochastic referent. It may be UPE to insure and have a fixed referent as well as not insure and have a stochastic referent; hence the necessity of the PPE refinement to select among such UPE choices. Koszegi and Rabin (2007) provide discussion of the relationship between insurance purchasing and referent expectations under the KR preference model.

arise in the form of apparent preference reversals depending on presentation. Interestingly, here the direction of preference reversals is predictable and grounded in a well-articulated theory of risky choice as opposed to an ex post psychological rationale (Hershey et al. 1982; Hershey and Schoemaker 1985). And, at least in the narrow context of this experiment, this model of choice is portable, providing predictive validity out of sample. Second, if referents can be manipulated via simple framing devices without physical endowments, then scope exists for marketers and policy makers to influence behavior with menus alone. These are hopeful directions for future research organizing inconsistencies in findings of risk aversion and assessing the extent of such inconsistencies in real-world behavior.

References

- Abdellaoui, Mohammed. 2000. "Parameter-Free Elicitation of Utility and Probability Weighting Functions." *Management Sci.* 46 (11): 1497–1512.
- Abdellaoui, Mohammed, and Han Bleichrodt. 2007. "Eliciting Gul's Theory of Disappointment Aversion by the Tradeoff Method." *J. Econ. Psychology* 28:631–45.
- Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *A.E.R.* 101 (2): 470–92.
- Allais, Maurice. 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine." *Econometrica* 21 (4): 503–46.
- Andreoni, James, and Charles Sprenger. 2011. "Uncertainty Equivalents: Testing the Limits of the Independence Axiom." Working Paper no. 17342, NBER, Cambridge, MA.
- Banks, James, and Zoe Oldfield. 2007. "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving." *Fiscal Studies* 28 (2): 143–70.
- Barberis, Nicholas, and Ming Huang. 2001. "Mental Accounting, Loss Aversion, and Individual Stock Returns." *J. Finance* 56 (4): 1247–92.
- Barberis, Nicholas, Ming Huang, and Tano Santos. 2001. "Prospect Theory and Asset Prices." *Q.J.E.* 116 (1): 1–53.
- Barberis, Nicholas, Ming Huang, and Richard H. Thaler. 2006. "Individual Preferences, Monetary Gambles, and Stock Market Participation." *A.E.R.* 96 (4): 1069–90.
- Becker, Gordon M., Morris H. Degroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Sci.* 9 (3): 226–32.
- Bell, David E. 1985. "Disappointment in Decision Making under Uncertainty." *Operations Res.* 33 (1): 1–27.
- Bleichrodt, Han, and Jose Luis Pinto. 2000. "A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis." *Management Sci.* 46 (11): 1485–96.
- Brookshire, David S., and Don L. Coursey. 1987. "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures." *A.E.R.* 77 (4): 554–66.
- Camerer, Colin F., Linda Babcock, George Loewenstein, and Richard H. Thaler. 1997. "Labor Supply of New York City Cabdrivers: One Day at a Time." *Q.J.E.* 112 (2): 407–41.

- Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Q.J.E.* 119 (3): 861–98.
- Card, David, and Gordon B. Dahl. 2011. "Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior." *Q.J.E.* 126 (1): 103–43.
- Castillo, Marco, and David Eil. 2014. "Taring the Multiple Price List: Imperceptive Preferences and the Reversing of the Common Ratio Effect." Working paper, George Mason Univ. http://mason.gmu.edu/~mcastil8/MCDE_52114.pdf.
- Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *A.E.R.* 96 (5): 1737–68.
- Costa-Gomes, Miguel A., Vincent P. Crawford, and Nagore Iriberri. 2009. "Comparing Models of Strategic Thinking in Van Huyck, Battalio, and Beil's Coordination Games." *J. European Econ. Assoc.* 7 (2–3): 365–76.
- Coursey, Don L., John L. Hovis, and William D. Schulze. 1987. "The Disparity between Willingness to Accept and Willingness to Pay Measures of Value." *Q.J.E.* 102 (3): 679–90.
- Crawford, Vincent P., and Nagore Iriberri. 2007. "Level- k Auctions: Can a Non-equilibrium Model of Rational Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica* 75 (6): 1721–70.
- Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *A.E.R.* 101 (5): 1912–32.
- Cubitt, Robin P., Chris Starmer, and Robert Sugden. 1998. "On the Validity of the Random Lottery Incentive System." *Experimental Econ.* 1:115–31.
- Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Juergen Schupp, and Gert G. Wagner. 2005. "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey." Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=807408.
- Epstein, Larry G., and Stanley E. Zin. 1990. "First-Order Risk Aversion and the Equity Premium Puzzle." *J. Monetary Econ.* 26:287–407.
- Ericson, Keith M. Marzilli, and Andreas Fuster. 2011. "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments." *Q.J.E.* 126 (4): 1879–1907.
- Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *J. Econ. Perspectives* 19 (4): 25–42.
- Gerardi, Kristopher, Lorenz Goette, and Stephan Meier. 2010. "Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data." Working Paper no. 2010-10, Fed. Reserve Bank Atlanta.
- Gill, David, and Victoria Prowse. 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *A.E.R.* 102 (1): 469–503.
- Goette, Lorenz, and Ernst Fehr. 2007. "Do Workers Work More When Wages Are High? Evidence from a Randomized Field Experiment." *A.E.R.* 97 (1): 298–317.
- Gonzalez, Richard, and George Wu. 1999. "On the Shape of the Probability Weighting Function." *Cognitive Psychology* 38:129–66.
- Grether, David M., and Charles R. Plott. 1979. "Economic Theory of Choice and the Preference Reversal Phenomenon." *A.E.R.* 69 (4): 623–38.
- Gul, Faruk. 1991. "A Theory of Disappointment Aversion." *Econometrica* 59 (3): 667–86.
- Haliassos, Michael, and Carol C. Bertaut. 1995. "Why Do So Few Hold Stocks?" *Econ. J.* 105:1110–29.

- Harbaugh, William T., Kate Krause, and Lise Vesterlund. 2001. "Are Adults Better Behaved than Children? Age, Experience and the Endowment Effect." *Econ. Letters* 70 (1): 175–81.
- Hardie, Bruce G. S., Eric J. Johnson, and Peter S. Fader. 1993. "Modeling Loss Aversion and Reference-Dependence Effects on Brand Choice." *Marketing Sci.* 12 (4): 378–94.
- Harrison, Glenn W., Morten I. Lau, Elisabet E. Rutstrom, and Melonie B. Williams. 2005. "Eliciting Risk and Time Preferences Using Field Experiments: Some Methodological Issues." In *Research in Experimental Economics*, vol. 10, *Field Experiments in Economics*, edited by Jeffrey Carpenter, Glenn W. Harrison, and John A. List. Greenwich, CT: JAI.
- Heffetz, Ori, and John A. List. 2014. "Is the Endowment Effect an Expectations Effect?" *J. European Econ. Assoc.* 12 (5): 1396–1422.
- Heidhues, Paul, and Botond Koszegi. 2008. "Competition and Price Variation When Consumers Are Loss Averse." *A.E.R.* 98 (4): 1245–68.
- Hershey, John C., Howard C. Kunreuther, and Paul J. H. Schoemaker. 1982. "Sources of Bias in Assessment Procedures for Utility Function." *Management Sci.* 28 (8): 936–54.
- Hershey, John C., and Paul J. H. Schoemaker. 1985. "Probability versus Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?" *Management Sci.* 31 (10): 1213–31.
- Herweg, Fabian, Daniel Muller, and Philipp Weinschenk. 2010. "Binary Payment Schemes: Moral Hazard and Loss Aversion." *A.E.R.* 100 (5): 2451–77.
- Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *A.E.R.* 92 (5): 1644–55.
- Isoni, Andrea, Graham Loomes, and Robert Sugden. 2011. "The Willingness to Pay—Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment." *A.E.R.* 101 (2): 991–1011.
- Jacobson, Sarah, and Ragan Petrie. 2009. "Learning from Mistakes: What Do Inconsistent Choices over Risk Tell Us?" *J. Risk and Uncertainty* 38 (2): 143–58.
- Kachelmeier, Steven J., and Mahamed Shehata. 1992. "Examining Risk Preferences under Higher Monetary Incentives: Experimental Evidence from the People's Republic of China." *A.E.R.* 82 (2): 1120–41.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *J.P.E.* 98 (6): 1325–48.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–91.
- Karle, Heiko, Georg Kirchsteiger, and Martin Peitz. 2015. "Loss Aversion and Consumption Choice: Theory and Experimental Evidence." *American Econ. J.: Microeconomics* 7 (2): 101–20.
- Knetsch, Jack L. 1989. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *A.E.R.* 79 (5): 1277–84.
- Knetsch, Jack L., and J. A. Sinden. 1984. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Q.J.E.* 99 (3): 507–21.
- Koszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Q.J.E.* 121 (4): 1133–65.
- . 2007. "Reference-Dependent Risk Attitudes." *A.E.R.* 97 (4): 1047–73.
- . 2009. "Reference-Dependent Consumption Plans." *A.E.R.* 99 (3): 909–36.

- Loomes, Graham, and Robert Sugden. 1986. "Disappointment and Dynamic Consistency in Choice under Uncertainty." *Rev. Econ. Studies* 53 (2): 271–82.
- Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education Programs." *Bus. Econ.* 42 (1): 35–44.
- Magat, Wesley A., W. Kip Viscusi, and Joel Huber. 1996. "A Reference Lottery Metric for Valuing Health." *Management Sci.* 42 (8): 1118–30.
- McCord, Mark, and Richard de Neufville. 1985. "Assessment Response Surface: Investigating Utility Dependence on Probability." *Theory and Decision* 18:263–85.
- . 1986. "'Lottery Equivalents': Reduction of the Certainty Effect Problem in Utility Assignment." *Management Sci.* 32 (1): 56–60.
- Meier, Stephan, and Charles Sprenger. 2010. "Present-Biased Preferences and Credit Card Borrowing." *American Econ. J.: Appl. Econ.* 2 (1): 193–210.
- Odean, Terrance. 1998. "Are Investors Reluctant to Realize Their Losses?" *J. Finance* 53 (5): 177–98.
- Oliver, Adam. 2005. "Testing the Internal Consistency of the Lottery Equivalents Method Using Health Outcomes." *Health Econ.* 14:149–59.
- . 2007. "A Qualitative Analysis of the Lottery Equivalents Method." *Econ. and Philosophy* 23:185–204.
- Plott, Charles R., and Kathryn Zeiler. 2005. "The Willingness to Pay—Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *A.E.R.* 95 (3): 530–45.
- . 2007. "Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?" *A.E.R.* 97 (4): 1449–66.
- . 2011. "The Willingness to Pay—Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Reply." *A.E.R.* 101 (2): 1012–28.
- Pope, Devin, and Maurice Schweitzer. 2011. "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes." *A.E.R.* 101 (1): 129–57.
- Post, Thierry, Martijn J. van den Assem, Guido Baltussen, and Richard H. Thaler. 2008. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show." *A.E.R.* 98 (1): 38–71.
- Prelec, Drazen. 1998. "The Probability Weighting Function." *Econometrica* 66 (3): 497–527.
- Rosch, Eleanor. 1974. "Cognitive Reference Points." *Cognitive Psychology* 7 (4): 532–47.
- Schmidt, Ulrich, Chris Starmer, and Robert Sugden. 2008. "Third-Generation Prospect Theory." *J. Risk and Uncertainty* 36:203–23.
- Schoemaker, Paul J. H. 1990. "Are Risk-Attitudes Related across Domains and Response Modes?" *Management Sci.* 36 (12): 1451–63.
- . 1993. "Determinants of Risk-Taking: Behavioral and Economic Views." *J. Risk and Uncertainty* 6:49–73.
- Schoemaker, Paul J. H., and John C. Hershey. 1992. "Utility Measurement: Signal, Noise, and Bias." *Org. Behavior and Human Decision Processes* 52:397–424.
- Segal, Uzi, and Avia Spivak. 1990. "First Order versus Second Order Risk Aversion." *J. Econ. Theory* 51 (1): 111–25.
- Shalev, Jonathan. 2000. "Loss Aversion Equilibrium." *Internat. J. Game Theory* 29:269–87.
- Smith, Alec. 2012. "Lagged Beliefs and Reference-Dependent Preferences." Working paper, Univ. Arizona. <http://www.u.arizona.edu/~alecs/papers/lbrdu.pdf>.

- Song, Changcheng. 2012. "An Experiment on Reference Points and Expectations." Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2580852.
- Starmer, Chris, and Robert Sugden. 1991. "Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation." *A.E.R.* 81 (4): 971–78.
- Stewart, Mark B. 1983. "On Least Squares Estimation When the Dependent Variable Is Grouped." *Rev. Econ. Studies* 50 (4): 737–53.
- Sydnor, Justin. 2010. "(Over)Insuring Modest Risks." *American Econ. J.: Appl. Econ.* 2:177–99.
- Thaler, Richard H. 1980. "Toward a Positive Theory of Consumer Choice." *J. Econ. Behavior and Org.* 1 (1): 39–60.
- Tversky, Amos, and Craig R. Fox. 1995. "Weighing Risk and Uncertainty." *Psychological Rev.* 102 (2): 269–83.
- Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *J. Risk and Uncertainty* 5 (4): 297–323.
- Wenner, Lukas M. 2015. "Expected Prices as Reference Points—Theory and Experiments." *European Econ. Rev.* 75:60–79.
- Wu, George, and Richard Gonzalez. 1996. "Curvature of the Probability Weighting Function." *Management Sci.* 42 (12): 1676–90.